

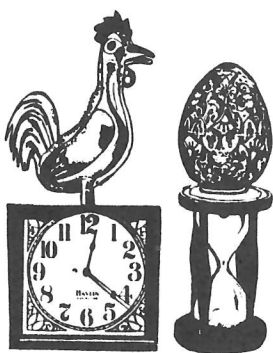
# TYPOLOGY RELATIONSHIP AND TIME

EDITED AND TRANSLATED BY  
VITALIJ V. SHEVOROSHKIN AND  
T. L. MARKEY





# TYPOLOGY RELATIONSHIP AND TIME





# Typology Relationship and Time

A Collection of Papers on Language Change and Relationship  
By Soviet Linguists

Edited and translated with a critical foreword by  
Vitalij V. Shevoroshkin and T. L. Markey



1986

KAROMA PUBLISHERS, INC.

ANN ARBOR

*Cover design by Lilian E. Stafford*

1986

Copyright © by Karoma Publishers, Inc.  
All Rights Reserved

ISBN: 0-89720-072-1 (Paper Only)  
Printed in the United States of America

## CONTENTS

FOREWORD, <i>T. L. Markey and Vitalij V. Shevoroshkin</i> . . . . .	vii
PROTO-LANGUAGES AS OBJECTS OF SCIENTIFIC DESCRIPTION. <i>V. V. Ivanov</i> (1980). . . . .	1
A PROBABILISTIC HYPOTHESIS CONCERNING THE OLDEST RELATIONSHIPS AMONG THE LANGUAGE FAMILIES OF NORTHERN EURASIA, <i>Aaron B. Dolgopolsky</i> (1964) . . . . .	27
REVIEW OF V. M. ILLIC-SVITYC'S DICTIONARY OF THE NOSTRATIC LANGUAGES (Vol. I), <i>V. V. Ivanov</i> (1972) . . .	51
REVIEW OF V. M. ILLIC-SVITYC'S DICTIONARY OF THE NOSTRATIC LANGUAGES (Vol. II), <i>V. V. Ivanov</i> (1977). . .	57
ON THE SO-CALLED "NOSTRATIC" LANGUAGES, <i>B. A. Serebrennikov</i> (1983) . . . . .	66
ON THE RECONSTRUCTION OF THE PROTO-INDO-EUROPEAN STOPS. GLOTTALIZED STOPS IN INDO-EUROPEAN, <i>T. V. Gamkrelidze and V. V. Ivanov</i> (1980). . . . .	87
BIBLIOGRAPHY . . . . .	109



## FOREWORD

**T. L. Markey and Vitalij V. Shevoroshkin**

-I-

Just how this little volume came into being certainly merits recounting. It is a strange, but not altogether implausible history these days that can make academic bedfellows of an Ohio farmboy and a former Soviet academician. The link between the two — and at the same time the impetus for the genesis of this volume, was curiosity, a "show-me-please" attitude. While Shevoroshkin had long been entirely convinced of the probative value of the Nostratic hypothesis, Markey certainly was not. Intrigued, he proposed that Shevoroshkin make a selection of what he considered the best of recent pro and con Soviet work on or about Nostratic and that the two then translate, edit, and preface the resulting collection for publication. That initial suggestion was made some four years ago now, and the project turned out to be much more arduous, much more time-consuming, and much more frustrating than either editor had assumed at the outset. Suffice it to say that East-West relations are not always of the easiest kind, and there have been "bureaucratic" delays. But all is well that ends well, and this little volume now appears, quite appropriately it seems to us, on the bicentennial of Sir William Jones's arresting announcement that Sanskrit bore a strong resemblance to Greek and Latin, so strong in fact that the resemblance could not be due to mere chance, but rather the result of genetic relationship. For his time, Jones's assertion was just as revolutionary as those of the proponents of Nostratic, who consent that Indo-European, Hamito-Semitic, Finno-Ugric, Dravidian, and South Caucasian (Kartvelian) are all ultimately related as "our" (Lat. *noster* → Nostratic) primordial mother tongue.

Even after the completion of this exercise, Markey

still retains a somewhat more skeptical attitude than does Shevoroshkin to what James A. Matisoff (p.c.) has called "megalo-comparisons", that is, attempts to demonstrate distant genetic relationships at great time-depths. There is, indeed, much that is highly controversial here. Then, too, this is also a highly controversial period in the history of more traditional comparative studies, a period that must, for example, now consider revamping, and extensively revamping at that, the usual phonemic inventory posited for Indo-European. As the contents of this volume show, Soviet work had a fundamental part to play in necessitating these revisions in what we have unquestioningly assumed to be the conventional wisdom. However, far from all of the assumptions that underlie Nostratic studies and its hypothesis are widely accepted. The Elamite-Dravidian theory has, for example, hardly received the wide approval Ivanov would like us to believe that it has. Then there are details, and one must read attentively in what still remains a somewhat jerky translationese despite numerous attempts at smoothing and polishing. For example, in the initial article by Ivanov the claim (pp. 7-8) that the second person singular ending *\*-tho* reflects PIE *\*-th<sub>2</sub>e* > (H)o in unstressed position, whereas it yields (H)a in stressed position. This is indeed possible, but would have tremendous consequences for the whole of Indo-European apophony. These consequences are never presented, nor are there any doubts expressed as Ivanov merely flits off, in his awesome gadfly fashion, to yet another topic.

However, both editors concur that the matter of providing information, of informing anglo-phone Westerners about the exciting possibilities raised by recent Soviet Nostratic scholarship, is far more important than assuming a self-righteously judgmental stance or offering a thumbs up or thumbs down vote on specific details. We have, after all, violated the original texts enough as it is. Far better, it seemed to us, to be informed than to remain provincial (and ignorant). But then, Nostratic inquiry, even labelled as such, is not all that recent, see, for example, Cuny (1943). Nevertheless, most Western scholars either are or have chosen to remain uninformed; unconvinced is quite another matter. In the light of

increasingly frequently accounts in the East (e.g. Čeyka and Lamprekht 1984), as well as such established classics as Diakonoff's (1965) remarkable Semito-Hamitic grammatical sketch and the critiques of Nostratic strewn about in such well-known historical surveys as Georgiev (1981 3rd ed.), it seems altogether odd that, in a textbook on universals and typology that also pays decent lip service to historical matters, Bernard Comrie (1981) never mentions Nostratic, and here is a scholar who seemingly prides himself on his knowledge of current trends in Soviet linguistics. But then that Nestor of American typological and universals research, Joseph H. Greenberg, also persistently overlooked Nostratic studies, or simply ignored them, even in his earlier (e.g. 1957) essays, which are hardly as programmatically biased as his later work. The suspicion then looms that Nostratic is a political matter, a figment of socialist imagination, a part of the party line. However, one fails to find Nostratic cited in historical surveys by Westerners (e.g. Haudry 1981) who are admittedly political in their presentations.

The suspicion that Nostratic is politically motivated is seductive. The reasoning goes as follows: wouldn't it make abundant good sense to relate all the languages of the Soviet Union, or at least the vast majority of them, to some single, albeit very distant, parent? This would be a sort of Marristic comparativism. Then the party line could go out that, although the various ethnic groups and linguistic entities in the Soviet Union might well appear under quite diversified garb today, never fear, for once they were all related. We would then have a new basis for socialist unity, but it would be that unfortunate kind of basis that was transformed into *ein Volk, ein Gott, ein Reich*.

Any allegation of political motives to account for the Soviet renaissance in Nostratic studies is quite false. This is an avenue of inquiry that actually started in the West with, for example, the findings of such scholars as Hermann Møller (b. 1850 - d. 1923), a Dane. Moreover, at the dawn of comparative studies there was everywhere a quest for common roots and meaningful insights into the origin of human language. Here one need but recall Adelung's 1781 classic, a work that appeared just five years before Sir William Jones made his startling

announcement about the putative affinity of Sanskrit, Greek, and Latin. Even more exciting, though certainly less well known, despite the fact that no less a historian of science than Michel Foucault pointed it out in his wildly popular *The Order of Things* (1971, pp. 233-4), is the fact that Russian nobility imported German intellectuals at least a decade prior to the appearance of Adelung's *Ursprung* for the purpose of gathering comparative linguistic data from the far reaches of the empire. Indeed, this collective enterprise was a priority for the Empirical Russian Academy of Sciences, see Bacmeister (1773), Guldenstadt (1787-1791), and note that the first of what was eventually a four volume set of language sketches appeared in St. Petersburg in 1787 with the title, *Glossarium comparativum totius orbis*, a work that also contains descriptions of no fewer than twenty-one Amerindian languages. The titles of these works, as well as copies of them, are all but totally unknown in the West yet today. The precise content of the intellectual life of 18th century Russian will remain vague until further efforts along the lines taken by Gary J. Marker (Department of History, SUNY, Stony Brook) to peer into the extent and nature of printing and publishing and the book inventories of elitist scholastic institutions and private individuals during the Empire are made.

If one were to magically telescope intellectual history in such a fashion as to omit the cleft between Empire and Soviet and simultaneously exclude Karl Marx & Cie. entirely, then Russian Nostratic studies would appear to continue an unbroken direction of inquiry that dates back to the 18th century, yet has its epistemic roots in the 17th century.

Boiled down the essentials, Soviet Nostratic studies rest on a classical ideal of knowledge, namely, on that logico-mathematical notion of understanding in which a coherent set of demonstrable propositions may be deduced from a small number of postulates. As Michel Serres (1982:71-83) pointed out in his essay in the origin of language, this is what the ideal of knowledge certainly was for Descartes, or Spinoza, or Leibniz, and recall that Leibniz (b. 1646 - d. 1716) holds the record for the largest corpus (albeit unpublished even today) of etymological writings of any of his contemporaries. Note just how

very mechanical Dolgopolsky's notion of system is in his contribution here. In Soviet hands, comparativism became a quasi-mathematical or logically manipulative system that seems virtually independent of the time variable. This is language history without history, but not quite. In fact, the time-depth required is simply too great for us to appreciate with any meaningful sociolinguistic detail. And if you were still hoping to catch a glimpse of surface phonetics in Indo-European studies, well you can forget it here. This is a brave new world of root structure, though time and again one finds oneself nodding approvingly as yet another piece of the giant etymological puzzle is correctly put in place.

Nostratic constructs are necessarily statutes in a mechanical system. Despite numerous objections, this is, it seems to me, a highly promising way to begin, but it remains just a beginning until or unless Nostratic constructs have some prognostic value conferred upon them. Here, we note that, in much of modern cognitive biology, the assumption of purely causal purpose has been replaced by the discovery of evolutive prognoses, that is, the uncovering of a rudimentary theory of preference. This is certainly what Rupert Riedl (e.g. 1979) has been all about so far. Apparently, Riedl would now evaluate linguistic structures not as a mechanical formulae, but as the best prognostic hypotheses available to a given system at a given time to ensure its survival. Be this as it may, the one undeniable saliency that emerges from this survey of Nostratic work is that the universalist and the typologist alike (and they have not always been united in the same person) can no longer overlook the possibility of an experimental linguistic history, while the converse is equally true: the historical linguist can no longer operate in a universal or typological vacuum. To do so would be to once again run the risk of committing such silly blunders as positing an entirely defective consonant system as the nuclear forebear for all that was to follow. The theoretician cannot do without history in the long run, and the historian must recognize theory, cf. Markey (1983, 1985).

The original versions of the articles translated here are, in their order of appearance:

- V. V. Ivanov. 1980. Prajazyki kak objekty opisaniya v izdanii "Jazyki mira". In: *Teoretičeskie usnovy klassifikacii jazykov mira*. Moscow: Nauka. Pp. 181-207.
- A. B. Dolgopolsky. 1964. Gipoteza drevnejšego rodstva jazykovyx semej Severnoj Eurazii s verojatnostnoj točki zvenija. *Voprosy jazykoznaniya* 2.53-63 (1964).
- V. V. Ivanov. 1972. *Ėtimologija 1972*. Moscow: Nauka, 1979. Pp. 179-84.
- V. V. Ivanov. 1977. *Ėtimologija 1977*. Moscow: Nauka, 1979. Pp. 179-84.
- B. A. Serebrennikov. 1982. Problema dostatočnosti osnovaniya v gipotezax, kasajuščix'sja genetičeskogo rodstva jazykov. In: *Teoretičeskie usnovy klassifikacii jazykov mira*. Moscow: Nauka. Pp. 6-62. (Translated portion here comprises pp. 47-62 only, a separate chapter which is entitled, *Nostratičeskie jazyki*).
- T. V. Gamkrelidze and V. V. Ivanov. 1980. Rekonstrukcija obščeeindoevropskogo jazyka. Glottalizovannye smyčnye v indoevropskom. *Voprosy jazykoznaniya* 4.21-35 (1980).

## -II-

The late Björn Collinder began his report to the conference on Indo-European linguistics at the University of California, Los Angeles (April 25-27, 1963), with the following words: "Holger Pedersen used to speak about a 'Nostratic' family of languages, comprising Indo-European, Semitic-Egyptian-Hamitic, Uralic, Altaic, Yukagir, and Eskimo. I do not think his thesis is proved. Can it be proved?" (Collinder 1966:199).

The following year, at yet another conference on Indo-European linguistics, but this time in Moscow, two scholars, V. M. Illič-Svityč and A. B. Dolgopolsky, both of whom had been working quite independently of one another, provided the first substantive evidence for the existence of a Nostratic macro-family of Eurasian languages comprising Indo-European, Afro-Asiatic/Hamito-Semitic, Uralic, Kartvelian, Dravidian, and Altaic. (In his earliest work, Dolgopolsky did not employ Dravidian data.) In their reports, they presented some partial results of their independently arrived at reconstructions of Proto-Nostratic. Arrestingly enough, their results coincided almost entirely (Illič-Svityč 1964, Dolgopolsky 1964a). In that same year, Dolgopolsky (1964b) published another seminal work with Nostratic reconstructions, and all of this appeared some two years prior to the publication of Collinder's statement as cited above.

Collinder was among the first to greet Illič-Svityč's and Dolgopolsky's findings with enthusiasm, an enthusiasm that, as Henrik Birnbaum was to put it later, bordered on the poetic. Subsequently, other established comparatists such as K. Menges, Poppe, Anttila, Ivanov, Dybo, Pisani, A. Lamprecht, M. Čejka, and P. Garde--to mention but a few--also accorded this outstanding piece of historical sleuthing high praise, see the summaries in Birnbaum (1977:56-9) and in Dybo's (1984) introduction to Illič-Svityč (p. 7, fn. 2), cf. Dybo (1978).

P. Garde insisted that Illič-Svityč's work provided abundant evidence of a rare versatility and solidity, and he pointed out that it included a considerable number of correspondences that were based on a highly rigorous method of phonological (or phonetological) comparison.

He concluded his review as follows: "It seems that we now have a real etymological dictionary, one that presents comparisons between language families, that, from now on, will be based on a solidly scientific basis."

The tragically early death of Illič-Svityč (b. 1934 - d. 1966), as well as Dolgopolsky's (1976) involuntary emigration to Israel, where he is now a professor at the University of Haifa, did not halt research on Nostratic in Russia. Former students of both Illič-Svityč and Dolgopolsky are currently at work in the Soviet Union in a highly productive fashion under the direction of V. A. Dybo in an external, extra-institutional seminar that bears Illič-Svityč's name. To date, members of this seminar have edited and published three parts of Illič-Svityč's Nostratic etymological dictionary: Vol. I (1971), Vol. II (1977), and, just this past year (1984), the first fascicle of Vol. III. They have also published extensively on Indo-European, Uralic, and Afro-Asiatic languages, as well as some important studies of non-Nostratic languages, again see Dybo's (1984) introduction to Illič-Svityč. The conference entitled "Nostratic Languages and Nostratic Linguistics," which was dedicated to the memory of Illič-Svityč and which was held in Moscow in 1977, proved to be highly successful: forty, mostly fairly lengthy, abstracts have now been published. Then, too, various topics from the area of Nostratic linguistics, as this new branch of historical inquiry has come to be known, have been presented at several subsequent conferences and gatherings in Moscow.

Despite the early success of Nostratic comparisons, based, as they are, on exacting phonological, grammatical, and semantic correspondences, this whole area of investigation is all but totally unknown in the West.

There are probably any number of reasons for this ignorance, some of which are entirely pragmatic, and some of which are certainly not. Among the purely pragmatic reasons is the fact that few Western scholars have an adequate reading knowledge of Russian to have ready access to the studies published to date, and then there are the particular political exigencies of the times which have made communication between Soviet and Western scholars both rare and difficult. But beyond--and surely more significant than--the purely pragmatic, there

is the present intellectual climate of the Western linguistic mainstream, specifically its static, ahistorical, and anti-processual bias, as well as its highly skeptical, often even openly hostile, attitude toward any kind of broader language comparison. This assessment is perhaps more valid for North America than for Europe, but be that as it may, the prominent result has been the virtual lack of any widely practiced cogent methodology for comparison or highly principled approach to reconstruction, particularly the reconstruction of proto-languages. Even when comparison has been attempted, such as in the realm of creolistics, the traditional trappings of diachrony have been remarkably absent: the majority of creolists remain blissfully ignorant of etymological pursuits. Then, too, few North American linguists now pass through the demanding mill of Indo-European studies: you can get a faster fix with clitic-climbing or pied-piping than with Lycian inscriptions or Hieroglyphic Luwian. These shortcomings may account for the numerous failures in recent attempts on the linguistic fringe to compare Eurasian and Amerindian languages, e. g. Penutian, Hokan, and so on, and, quite recently, comparisons between Indo-European and Afro-Asiatic languages. The repeated lack of success here has, as might be expected, led to the categorical conclusion that any broad-scale comparison between seemingly (on the surface) disparate languages is ultimately doomed to failure, an adventure into the realm of whimsy. Further fall-out from this anti-historical, anti-comparative climate may also account for the fact that Soviet publications in Nostratic from the past two decades remain unused or uncited. But it is not only Soviet scholarship in this area of remote linguistic affiliations that has been consigned to a conspiracy of silence, for many very revealing comparisons (albeit largely without proper reconstructions) on a broader scale by an American linguist as distinguished as Joseph H. Greenberg, who built on the previous musings of Carl Meinhof, remain unnoticed, relegated to the great dustbin of linguistic history.

The present collection of articles is a first step toward bridging a gap and toward drawing the attention of Western linguists to a whole range of rather remarkable accomplishments by Russian scholars, not only in the

area of distant genetic affiliations between and among languages, but also in the more theoretical arena of designing the proper principles and techniques for comparison *per se*. In making our selections, we have obviously concentrated on Nostratic; this is the central topic of the papers by Dolgopolsky, Ivanov, and Serebrennikov; but the methodology of comparison and the principles for reconstructing proto-languages are also discussed by both Dolgopolsky and Ivanov, while Ivanov and Gamkrelidze present recent data in support of their new and highly controversial theory of Indo-European consonantism. Here, then, one will find attempts to define macro-families (phyla) and to reconstruct proto-languages, *plural*, not just Nostratic; and though many of these papers are indeed centered around the Nostratic hypothesis, not all are unreservedly pro Nostratic: Serebrennikov's paper is distinctly critical. Thus, we have tried to be objective in our inaugural presentation.

Nevertheless, those pragmatic factors alluded to above did hamper us. It is in part due to them that we were unable to include several other, equally significant, contributions. For example, O. Mudrak's paper on the Nostratic character of Eskimo-Aleut has, unfortunately, been "lost" in the mail; and, for a number of similarly tedious reasons, S. Starostin's report on the Nostratic character of Niger-Kordofanian was unobtainable. Then, too, by the time we received S. Starostin's article on the reconstruction of Proto-North-Caucasian and Proto-Yeniseyan, a paper in which these languages are compared with Proto-Sino-Tibetan within the framework of a putative Sino-Caucasian macro-family, it was simply too late, see Starostin (1982), and note that S. Nikolaev's paper on Dene-Caucasian (he later added Na-Dene languages to Sino-Caucasian, thereby significantly broadening the horizons of this macro-family, which is now termed Dene-Caucasian) is still in print; see the further reports on this work by Shevoroshkin (1981), Ivanov (1983:154-6), and in Dybo's (1984:5) introduction to Illič-Svityč. Nikolaev's paper on a possible North-Caucasian substratum for the languages of Europe, as well as S. Čirikba's stimulating article on the North Caucasian character of Basque, are also in this collection.

Tempting as it may have been, we finally decided

against publishing lists of reconstructions, as they would have been primarily interesting for the narrow specialist alone: a list of some two thousand stems reconstructed by Starostin and Nikolaev for Proto-North-Caucasian, as well as a list of some 700 Proto-Hokan roots only recently reconstructed by D. Leščiner.<sup>1</sup> We hope that these and similar studies will soon be made available for Western scholars.

### On the Methods of Comparison

Ivanov's article, "Proto-Languages . . .," and Dolgopolsky's article, "Hypothesis . . .," provide some profound insights into both the techniques for comparison and reconstruction and the present state of comparative research. D. describes the initial step in comparison where the goal is to discover whether or not certain languages are even remotely related. It is possible to determine both the varying degrees of consonantal stability and the regular patterns of evolution for certain types of consonants. D. underscores the fact that it is imperative to work with many languages simultaneously, as this provides many more possibilities for reciprocal evaluation and checking. Comparisons based on but two languages seldom provide secure results, and it is hardly by chance that binary comparisons, ever popular in this country, have generally failed. Phonetic/phonetological correspondences must be worked out in the process of comparison; it is altogether wrong to make comparisons first and then compile an inventory of correspondences. It practically goes without saying that all kinds of correspondences must, in the final analysis, be fully explicit. For example, each phoneme in a given radical in Language A has to match the respective phonemes in what is presumably the corresponding radical in Language B; semantic correspondences must be equally explicit; and so on.

One of the most important points in D.'s method is his determination of a hierarchy of relative stability for basic lexemes. Such meanings are represented by words that are rarely, or never, subject to replacement by other words (with the same meaning). Then, too, such words are rarely, or never, subject to borrowing. D. gives the hierarchy for the fifteen stablest lexemes: 1. *I/me*, 2. *two/pair*, 3. *thou/thee*, 4. *who/what?*, 5. *tongue*, 6. *name*,

7. *eye*, 8. *heart*, 9. *tooth*, 10. *no/not* (negative prohibitive), 11. *finger nail/toe nail*, 12. *louse*, 13. *tear*, 14. *water*, 15. *dead*. Obviously, this list includes both grammatical and lexical items. On the basis of D.'s asides and notes, one can easily compile a list of the ten next stablest lexemes. 16. *nit*, 17. *moon*, 18. *hand*, 19. *night*, 20. *blood*, 21. *horn*, 22. *full*, 23. *sun*, 24. *ear*, 25. *salt*. This inventory of twenty five items/semantic values is mainly comprised of nouns, as a rule those for body parts, only a few adjectives, and no verbs, though if one were to continue, verbs would soon appear.<sup>2</sup>

The development of this list is a major step toward resolving the debate over whether or not the Swadesh list of basic lexical items is a legitimate tool for comparative purposes and what, in fact, must be included in the first one hundred items. From D.'s article it is clear that we need such a list and that we no longer have to argue about what to include: the list is a hierarchy, and the numbers indicate the relative position of a given meaning in this hierarchy. Only one thing is left to define: how many meanings are required for a given comparison? The list will come automatically, and it goes without saying that more than one word will sometimes fit a given meaning.

Ivanov's article stresses the convenience of defining languages that belong to a certain (macro-) family well prior to beginning the actual work of reconstructing the respective proto-language, cf., on this matter, the numerous well-known works by Joseph H. Greenberg. The results of this procedure, which are achieved from comparisons of the stablest morphemes, are generally confirmed by subsequent reconstruction in what is a circularity of processing that is both heuristically and probatively effective. One must bear in mind, however, that the existence of a certain genetic affinity between languages does not necessarily imply that a respective proto-language has, indeed, ever existed, but may, in fact, be purely hypothetical, a logical construct spoken by no one. In certain cases, there is simply no time left for the existence of a proto-language. One of Ivanov's examples of such a case, namely, that of "Proto-Altaic," is especially enlightening. Despite the reconstruction of Proto-Altaic by Illic-Svityč, it is now clear that Ivanov, Dolgopolsky,

and various other scholars prefer to speak of three proto-languages (i.e. Turkic, Tungusian, and Mongolian) as having originated directly from Proto-Nostratic. Nevertheless, this does not exclude assuming the notion of an Altaic "family" of languages that includes Turkic, Tungusian, and Mongolian, as well as Japanese and Korean. Indeed, if one analyzes Illič-Svityč's Proto-Altaic attentively, one realizes that it does not differ greatly from Proto-Nostratic; only Altaic among the other East-Nostratic languages (i.e., Altaic, Uralic, and Dravidian) has three types of stops, whereas all the others have but one type; Altaic vowels are practically identical with those posited for Nostratic, etc. In various studies by the late V. Cincius (e.g. 1984:18), an additional type of Altaic affricate was added to Illič-Svityč's two types, so that, in the improved version, Altaic now has three types of affricates vs. one type in both Uralic and Dravidian; again, Altaic affricate become virtually identical with those posited for Nostratic. Dolgopolsky (1972) has pointed out that revised Turkic, Tungusian, and Mongolian data show differences between the systems of affricates and fricatives in these three proto-languages that once again exclude reconstruction of a supra-dialectal Proto-Altaic. This substantiates Cincius' claim that the Proto-Altaic affricates are, indeed, those of Proto-Nostratic.

### **On Dravidian**

We have included the "anti-Nostratic" article by Serebrennikov and footnoted it with editorial comments which show that S.'s critical remarks are invalid.<sup>3</sup> In the same collection of articles as that in which S.'s article appeared, there is a paper by Andronov (1982) that has (Pp. 186-90) some brief remarks on Nostratic. During a visit to Ann Arbor in August of 1983, Dolgopolsky kindly agreed to comment on Andronov's remarks:

"Andronov's notes do not refer to the Dravidian cognates given in the main part of Illič-Svityč's Nostratic etymological dictionary, but rather to those in the introductory part of the first volume of the dictionary, see Illič-Svityč (1971:3-37). This introductory section contains tables of comparisons that have only a preliminary character; their sole purpose is to validate the preliminary possibility of genetic comparisons between Dravidian, Uralic, Altaic, Indo-European, Afro-Asiatic, and Kartvelian.

These preliminary correspondences were later checked over by Illic-Svityč, and only part of them were included in the dictionary proper. If one were to go back to these tables of preliminary correspondences now, one would quite naturally emend some of them. It is unclear why Andronov concentrated on this preliminary part, but totally ignored the main corpus of the dictionary.

Andronov (1982:188) states that: "The vast majority of Dravidian correspondences with Nostratic pronominal stems as far as case, number, and verbal formants are concerned can be explained as entirely fortuitous similarity: etymologically they have nothing in common." Now, of course, 'vast majority' does not mean 'all', so there is a minor residue that cannot be explained by appealing to fortuitous coincidence. Therefore, we are on firm etymological grounds. But this is possible only if there is an underlying genetic relationship."

Upon closer scrutiny, it turns out that Andronov's assertion that 'a vast majority' of correspondences are etymologically unrelated is unfounded. Let us take his own example. Drav. *\*tā-m* (where *-m* = plural) means not only 'themselves' but also (and primarily) 'they', cf. Burrow-Emeneau (1961:No.2582): "Tamil *tām* 'they, themselves', Telugu *tāmu* . . . 'they, themselves', Naiki, Parji *tām* 'they, themselves', Goda *tammā* id., Kurukh *tām* id.," etc. It is highly unlikely that 'they' is a late development of 'themselves', especially in view of the fact that 'they' is present in all branches of the Dravidian family. It is far more likely that 'they' is the archaic meaning of this form, and hence, that it is open to licit comparison with (demonstrative) pronouns with comparable shapes in other Nostratic languages.

Andronov's assertion that Drav. *\*nā-m* 'we' derive from *\*yām* (p. 188) is an entirely personal opinion and contradicts probative etymologies accepted by all other Dravidianists. According to K. Zvelebil (1977:378), Drav. *\*nām* 'we' (1st pers. pl. inclusive) differed from *\*yām* 'we' (1st pers. pl. exclusive). The inclusive/exclusive-distinction is preserved in many Dravidian languages, e.g. Old Tamil *nām* 'we' (inclusive) vs. *yām* 'we' (exclusive). Therefore, comparison of Drav. *\*nām* 'we' with IE *\*ne-*, *\*nō* 'we', Kār tv. *\*naj* 'we', and Afro-Asiatic *\*naḥnu*, *\*n-* 'we' is possible.

Andronov's arguments against purely lexical comparisons

are mainly both subjective and unsubstantiated. Thus, for example, even if Drav. *\*ul-* actually had a value 'to be within' (so Andronov 1982:188), although its basic meaning was on the order of 'to be', it is still comparable with Uralic *\*ol-* 'to be', Turkic *\*ol-/bol-* id., Afro-Asiatic *\*wʔl*, as these forms have semantic values very close to 'to be (within)'. Phonologically, too, this comparison is satisfactory: the initial Nostratic sonorant was *\*w-*, see Illič-Svityč (1971:150 and the table there, sub *\*w-*). Even if Andronov's remarks concerning the five roots he mentions were correct, they would still not refute the entire corpus of Nostratic correspondences, but merely supply some additional specification. Andronov's line of reasoning is about on a par with, say, the contention that *a*-vocalism in North Germanic *\*alu-* 'beer, ale' indicates that North Germanic is non-Indo-European.

Here one could add that, in his introduction to Illič-Svityč's dictionary, Dybo (1984) stressed the fact that Andronov's thesis about Drav. *\*nam* < *\*yām* 'we' is based on an otherwise unattested shift of *\*y* > *\*n*. Then, too, a close phonetological similarity between *\*nām* and *\*yām* is an illusion: *-m* is a plural suffix employed in several other forms; *\*ā*-vocalism figures in several other pronominals with deictic force and may well have been a primordial sign of directionality in conjunction with other elements, cf. *tā-m* 'they, themselves'. What remains is *-n* vs. *y-*, two quite different elements without a derivational relationship. Dybo (1984:10, fn. 9) also remarks that Drav. *\*pēr-* 'big, great' has, contrary to what Andronov says, nothing to do with Nostr. *\*berg[i]* 'high', which Illič-Svityč (1971:177, cf. p. 35 *\*per*) compares with a nearly homonymous, but different, Dravidian root, namely, *\*pēr* 'high'.

Dybo laments the fact that the critical remarks vis-à-vis Nostratic by Serebrennikov and Andronov are neither constructive nor correct. He points out that valid criticism implicates a clear-cut distinction between, on the one hand, scientific facts and, on the other hand, mere conjecture divorced from scientific facts. Second guessing is not science. In the foreword to the collection that contains Andronov's article, we find the following assessment: "Andronov writes disapprovingly about Nostratic attempts to include Dravidian in the Nostratic

family, and he does point out some factual errors." Dybo then goes on to show that the so-called "factual" mistakes are in reality highly idiosyncratic appraisals on the part of Andronov or mere conjectures that find no support whatsoever in the accepted linguistic data, nor, for that matter, any support from recognized Dravidianists. Dybo also emphasizes the fact that, despite substantial progress subsequently in comparative historical linguistics and repeated attempts to discover errors, virtually no mistakes were found in the Nostratic comparisons that Illič-Svityč had made over two decades ago. This in itself is a strong vindication of both the theory and the precise prognostic validity of Illič-Svityč's research, but this does not mean that his work was flawless, and Illič-Svityč himself always welcomed constructive criticism.

Dybo (1984:7) also characterizes some previous anti-Nostratic criticism as being "an emotionally negative reaction to what is fundamentally a new direction in research." We know of but one published instance of such emotionally charged criticism, namely that by G. Doerfer in his *Lautgesetz und Zufall* (Innsbruck, 1973). Extracts from Doerfer's conclusion are illustrative. "Schluss mit dem mystisch-verschwommenen Omnicomparativismus! Historische Sprachwissenschaft: Ja! Rekonstruierende Sprachwissenschaft: Ja! Glottogonische Sprachwissenschaft! Nein!" (p. 122). We have no doubt but what Illič-Svityč would have stamped his approval on this statement, for by all indications he was one of the most careful, meticulous, and accurate representatives of the long tradition of comparative historical linguistics. He even pointed out numerous errors on the part of "omnicomparatists." So we are here confronted by emotion rather than logic and believe that few scholars would agree with Doerfer's (1973:112) remark that: "In Wirklichkeit schafft ja Illič-Svityč gar keine Pionierarbeit," or (p. 113) that: "Ficks Material [from his Indo-European dictionary of 1868] ist bereits reicher und zuverlässiger als dasjenige von Illič-Svityč." We quite agree, however, that Doerfer was sincere in his opinion and wished to distance himself from Nostratic theory and its empirical support. It appears that he was so ashamed of using "unreliable" material from Illič-Svityč's dictionary in his own article, "Basic Lexicon and the Altaic Problem"

(in Russian in *Vosprosy jazykoznanija*, Vol. 4 (1981), p. 35 ff.), that he never indicated that Illič-Svityč was the source of his data. In his *Lautgesetz und Zufall* (1973:113), he states that all the laws of probability speak against Illič-Svityč's findings; and he then launches a vicious attack against Dolgopolsky's 1964 article (in this collection); but this is an article that seeks to demonstrate that the laws of logical probability validate the thesis of distant linguistic relationships. Incorrectly, too, Doerfer asserts that Dolgopolsky tried to prove a genetic relationship between Indo-European, Semitic, Uralic, Altaic, Kartvelian, Chukchee-Kamchatkan, and Sumerian. In comparing these languages, however, Dolgopolsky showed rather that Sumerian is not related to any of the other languages in this list. Moreover, Doerfer reproaches Dolgopolsky for not postulating phonetic laws; but it was one of Dolgopolsky's goals in this article to show how to begin comparative research, and as a first step Dolgopolsky set up ten types of underlyingly basic consonants, a very attractive start indeed. Some items compared on this level of abstraction would perhaps be discredited and later dumped, and Doerfer is thus unjustified in criticizing these comparisons as a final version of what Dolgopolsky had fully intended as highly preliminary and very rough approximative comparisons of a few lexical items. Dolgopolsky was here exposing the nuts and bolts of the mechanics of developing a comparative methodology, not the final, fully refined product of that methodology. This is on a par with criticizing the Standard Theory of generative grammar for failing to accomplish what the Revised Extended Standard Theory has accomplished. But perhaps Doerfer's sorriest misunderstanding of Dolgopolsky's intentions is revealed by the following statement in his *Lautgesetz und Zufall* (1973:114): "Dolgopolsky hat willkürlich 15 Wörter herausgegriffen, wo sich bei den verglichenen Sprachfamilien relativ viele Ähnlichkeiten fanden." No, Dolgopolsky did not pull out the fifteen words arbitrarily. He inventoried 15 meanings, not 15 words, and he did so after having surveyed some two hundred languages. The purpose of his survey was to discover the relative degrees of stability for the most stable basic lexemes of essential lexico-semantic values in the world's languages. He wanted to show that the

lexical items that represent these meanings are seldom (or never) replaced by other words with the same meaning in any given language—a highly significant discovery, by the way. Doerfer then incorrectly asserts that Dolgopolsky had not taken mere fortuitous coincidence or the effects of purely symbolic phonetics (sound symbolism) into account, but the whole idea behind Dolgopolsky's paper was to demonstrate that, in these highly specific cases of stability, the effects of sound symbolism and mere coincidence are ruled out, thereby yielding a basis for determining a genetic relationship, however distant that relationship might appear on the surface.

### **On Indo-European**

As has always been the case since the dawn of modern historical linguistics in the latter portion of the 18th century, the vast majority of comparatists have specialized in the study of Indo-European. As Indo-European is also considered a member of the Nostratic macro-family, it might prove interesting to a major segment of the community of comparatists to concentrate here on two topics: (1) just how reliable, from a Nostratic point of view, are the new reconstructions of the Proto-Indo-European sound system? And (2), quite apart from the broader scope of Nostratic comparison, how reliable are the recent binary comparisons of Indo-European and Afro-Asiatic?

As far as the first question is concerned, we shall briefly discuss two issues: (1) the new reconstruction of the Indo-European system of stops (see the paper by Gamkrelidze and Ivanov in this collection) and (2) the new reconstructions of Indo-European laryngeals (and the closely related matter of the origins of Indo-European vocalism).

We quite agree with G. and I.'s contention that the traditionally reconstructed system of Indo-European stops of the type *T*, *D*, *DH* (voiceless, voiced, and voiced affricate) is in need of repair, or, at the very least, re-interpretation. It was, it now seems, rather uncritically modeled on that found in Sanskrit; and, as Roman Jakobson pointed out, it has no typological support. It is quite probable that the Germanic (*Th*, *T*, *\*Dh*), Anatolian (Hittite-Luwian) (probably the same as that posited for

Germanic), and Armenian (*Th*, *T*, *D*) systems are more archaic than any other systems of stops in the Indo-European dialects. This thesis finds support in the fact that Anatolian (Hittite-Luwian) was one of the first *speech communities* to break away from the Indo-European *language community*. These Anatolian languages lack the feminine, which was a later innovation in the remaining Indo-European dialects. Hittite-Luwian had but two "genders," common and neuter, which probably originally reflected a contrast between animate and inanimate respectively, or, perhaps reasoning vs. unreasoning (as in some Caucasian languages), or some such classifier system later grammaticalized as gender. Returning to the phonological problem at hand, it is thus highly likely that either *Th*, *T*, *Dh* or *Th*, *T*, *D* reflected the primordial Indo-European state of affairs. In languages in which *Th* became "weakened" to *T*, the original *T* became *D*. This shift presupposes that *D*, but not *Dh*, was originally the third member of the triad, cf. conceivable alternative shifts: *\*Th* > *T*, *\*T* > *D*, *\*D* > *Dh*. This view is supported by Nostratic, cf. Nostr. *\*T'*, *\*T*, *\*D*; the resulting shift, *\*Th*, *\*T*, *\*D*, was original both for Indo-European and the most archaic Altaic proto-languages (= Proto-Altaic *\*Th*, *\*T*, *\*D*, according to Illic-Svityč).

It is possible, though highly speculative, that the shift from Nostr. *\*T'* to Proto-Indo-European and Proto-Altaic *\*Th* was connected with the marked reduction (and later elimination in most cases) of an originally developed system of glottal, laryngeal, and pharyngeal phonemes (Nostr. *\*ʔ*, *\*h*, *\*h̥*, *\*ʕ*), cf. infra. In line with recent findings in general phonetics (see the excellent survey by John C. Catford in his Fundamental Problems in Phonetics [1977]), we can now interpret the Proto-Indo-European system of stops (types *Th*, *T*, *D*) in the following way:

- I.) *Th* Intensive, aspirated (cf. Modern Icelandic, etc.), with relatively large glottal opening during stricture (another possibility:  $\bar{T}$  as intensive, with long duration, with a further development of  $\bar{T}$  to *Th* in Germanic, Anatolian, and Armenian, but weakening of  $\bar{T}$  to *T* in the other dialects, a less likely, but still possible scenario ).

- II.) *T*      Non-intensive, with narrowed glottal opening, potentially voiced, but actually voiceless (cf. voiceless *b*, *d*, *g* in some modern German dialects, in some English forms, etc.).
  
- III.) *D*      Non-intensive, with relaxed, vibrating vocal chords; considerable amounts of air escape through a chink, generating whisper (either voiced or unvoiced phonetically; in the latter case, Proto-Italic *\*th* and Gk. *th* can be explained as originating directly from the IE *D*-type dental).

Our second element, *T*, in this triad did not change in phonetically/phonetologically archaic languages (Germanic/Anatolian/Armenian); but it became voiced, as a result of systemic pressure, in those languages in which IE *\*Th* was "weakened" to *\*T*, that is, in all other Indo-European dialects. This was a simple change, as IE *T*-type stops were potentially voiced.

Our third element, *D*, was phonetically/phonetologically very close to voiced aspirates (which are produced with relatively large glottal opening during the stricture and with large amounts of escaping air that generates whispery voice); it was unstable, and a developmental scenario with *D* > *dh* > *d̥* was highly probable, cf. the background events required for Germanic and Anatolian and note Proto-Indo-Iranian *\*dh*. On the other hand, as *D* was phonetically/phonetologically very close to lax [*d*] as well, we find *d* in Armenian, and the merger of *T* and *D* dentals resulted in *\*d* in both Proto-Baltic and Proto-Slavic. Some Anatolian languages appear to display [θ], cf. Italic.

Only IE *\*T* was pronounced with a narrowed glottal opening and, hence, no [th] or [dh] or the like resulted in Indo-European dialects; this accounts for the lack of aspirated stops. On the other hand, both intensive *\*Th* and "weak" *\*D* were pronounced with a large glottal opening and with the escape of large amounts of air, and this resulted in aspiration in many Indo-European dialects.

The third element, *D*, was the one element that was closest to a voiced stop. This assumption is supported

by the fact that virtually all loans from Semitic to Indo-European show (or showed) *D* in both Semitic and Indo-European. This evidence also supports the claim that the second element, *T*, was actually phonetically/phonotologically voiceless, but not intensive, cf. our definition: non-intensive, phonetically/phonotologically voiced, though actually voiceless. Hence, some very few Semitic forms with voiced stops were reflected with *T* in Indo-European.

Now, if we compare our results with the new reconstruction by Gamkrelidze, Ivanov, and Hopper, we see a principal difference in the interpretation of the second element of the triad, but a new coincidence in the first and third elements:

	Gamkrelidze - Ivanov	Our Reconstruction
1.	<i>T(h)</i>	<i>Th</i>
2.	<i>T'</i>	<i>T</i>
3.	<i>D(h)</i>	<i>D (&gt; D, Dh)</i>

We realize that typological considerations compelled Gamkrelidze, Ivanov, and Hopper to reconstruct the second element as a glottal; and even in our reconstruction it is the one element with a narrowed glottal opening; but here the resemblance ends. The Indo-European system of stops was a "weakened" system when compared with that posited for Nostratic; the shift from Nostr. *T* (second element) to its counterpart in Proto-Indo-European allows no place for glottalization. Nostratic glottal stops resulted in IE *Th* (Gmc. *\*th*, Gk. *t*, etc.) and not in *T*. Failure to appreciate this realization accounts, in part, for the lack of success in a recent attempt to compare (developmentally, genetically) Indo-European and Afro-Asiatic (Semitic-Hamitic/Hamito-Semitic) in terms of the new Gamkrelidze-Ivanov-Hopper reconstruction, see *infra*.

Here it should be pointed out that, to date, despite this new theory's challenge to the standard, traditional results of Indo-European comparison, no entirely persuasive arguments have yet been made in favor of either the old or the new (Gamkrelidze-Ivanov-Hopper) hypothesis. Obviously, the underlyingly genetic epistemology on which

our reconstruction is based--and this also applies to the traditional view and that by Gamkrelidze-Ivanov-Hopper --is not strictly scientific, since no real testing is possible. Nevertheless, however briefly, we have outlined typologically based quantification of the restrictions involved and offered formally definable phonetic/phonetological implications, two essential requirements for any valid thesis of historical sound change. The establishment of a broadly based evaluation procedure must be left for future inquiry; a precise definition of the relationship between (typological) naturalness and grammatical simplification in the course of linguistic change has yet to be supplied. If, for example, SVO is regarded as the typologically natural order (where object-first ordering is considered the least natural, most severely complicated, and empirically least common ordering), why is it that simplified languages, e.g., all true creoles, display SVO, in fact the only result observed to date? Specification of the contents and dynamics of the relationship between typological naturalness and simplification on all grammatical levels, phonological as well as syntactic, is one of the most pressing requirements for dramatically insightful historical interpretations. Obviously, we have once again argued in favor of a retention of the voiceless/voiced/voiced aspirate contrast of Indo-European in Anatolian, and, for the voiced/voiceless contrast at least; and our view has not been successfully opposed to date. Further substantiation of the full range of contrasts, particularly in the Anatolian group, is another matter--and one that must be left for future inquiry.

Let us now turn to Indo-European laryngeals and vowels. Here, too, the findings of Nostratic theory place some well considered constraints on theories that are based on data from Indo-European alone. We may take a further step here and not only use the comparisons of Indo-European with the other Nostratic languages employed by Illič-Svityč and Dolgopolsky some two decades ago, but also introduce further comparisons; and on this point we are pleased to state that our results coincide with those independently achieved by Dolgopolsky in his recent revision of Indo-European reconstructions.

According to Classical Nostratic Theory, which was

based on what has been taken as the standard, classical reconstruction of laryngeals for Proto-Indo-European, IE  $*\hat{h}$  (palatalized  $h$ ),  $*h$ , and  $*h^w$  originate from Nostratic postvelars  $*q$  and  $*g$ , laryngeal, pharyngeals, and glottal stop  $*ʔ$ . All merged in Indo-European: three types of IE  $h$ -fricatives resulted when, supposedly, Nostratic vowels merged as Indo-European  $*e$ . The quality of IE laryngeals depended on the quality of the original neighboring vowels: whence, IE  $*\hat{h}$  in words in which the initial consonant of the Nostratic protoforms was followed by Nostratic front vowels, viz.  $*h$  before  $*a$ , and  $*h^w$  in words where the following Nostratic vowel was  $*o$ ,  $*u$ , or  $*\ddot{u}$ . A similar explanation of the three types of IE gutturals ( $\hat{K}$ ,  $K$ ,  $K^w$ ) proved successful (cf. infra); but in the case of the laryngeals nothing could be proved, and several contradictions and shortcomings soon became obvious.

The Anatolian data clearly indicated that, in Indo-European, there were at least two different laryngeals that could appear before the same Indo-European vowel. One IE laryngeal resulted in Anatolian  $*h$  (> Hitt., Luw.  $\hat{h}$ -), while the other disappeared in Anatolian. To explain, say, Nostr.  $*qant$ - 'front' as having first resulted in IE  $*hent$ - (Nostr.  $*t'$  regularly developed as IE  $*t$ ; Nostr.  $*a$  indicates that the IE laryngeal was  $*h$  and not  $*\hat{h}$  or  $*h^w$ ), which then became  $*ant$ - in the various IE dialects, seemed highly artificial. Why, indeed, must one reconstruct Nostr.  $*a$  > PIE  $*e$  >  $*a$  (in most individual proto-dialects of Indo-European), rather than Nostr.  $*a$  > IE  $*a$ , when all dialects reflect  $a$  and not  $e$ ? The answer is obvious. Classical Nostratic Theory required that all Nostratic vowels first become  $*e$ ,  $*ie$ ,  $*we$  in Indo-European. After Indo-European laryngeals, however, the quality of the three types of Nostratic vowels seemed to be preserved. Vowels  $*\ddot{a}$ ,  $*e$  resulted in IE  $*e$ ;  $*i$  became  $*e$  or  $*ai$ / $*\hat{i}a/i$  under certain conditions,  $*a$  remained as  $*a$ , but labial vowels resulted in  $*o$  or  $*au$ / $*wa$ / $*u$  this was an innovation. A way around this was to reconstruct at least two IE laryngeals: "strong"  $*X$ , which resulted in Hittite and Luwian  $h$ , and "weak"  $H$ , which disappeared in Anatolian; but their quality was independent of that of neighboring vowels and vice versa.

Here, too, Nostratic theory revealed its principal

correctness, as evidence from the Nostratic languages provided strong support for this reconstruction. Those languages, namely Afro-Asiatic and Kartvelian, in which the original system of Nostratic obstruents had been preserved well enough showed that Nostr. *\*q*, *\*g*, *\*x*, *\*ʔ* (postvelar stops and fricatives, all strong) became IE *\*X* (> Hitt., Luw. *ḫ*), while the "weak" Nostratic series (*\*ʔ*, *\*q̣*, *\*ḥ*, *\*h*) became *\*H*. They disappeared in Hittite in initial position, but lengthened *C* or *V* in non-initial position, and this was assumed to be the situation in other archaic Indo-European languages as well.

As for vowels which followed (or preceded) the laryngeal, postulation of the following changes was supported by data from the East Nostratic languages (Uralic, Dravidian, Altaic) which apparently retained the original Nostratic system quite well:

Nostr. <i>*e</i> , <i>*ḁ</i>	>	IE <i>*e</i>
Nostr. <i>*i</i>	>	IE <i>*e</i> or <i>*ai/*ia</i> (> <i>*ie</i> )
Nostr. <i>*a</i>	>	IE <i>*a</i>
Nostr. <i>*o</i>	>	IE <i>*o</i>
Nostr. <i>*u</i> , <i>*u̥</i>	>	IE <i>*o</i> or <i>*au/*wa</i> (> <i>*we</i> )

By way of illustration, we shall now provide a few examples to demonstrate these correspondences. (An article on the origin of IE laryngeals and vowels in the neighborhood of laryngeals by M. Kaiser and V. Shevoroshkin will soon appear in *Journal of Indo-European Studies*, and an article on the origin of IE vowels by Dolgopolsky will also soon appear.)

Examples with an IE "weak" laryngeal < Nostr. *\*ʔ*, *\*q̣*, *\*ḥ*, *\*h*:

i. Nostr. *\*ʔesA* 'dwell, stay' > IE *\*Hes-* (cf. Hitt. *es-* 'to be'); for *\*ʔ*, cf. Afro-Asiatic *\*ʔjṣ/č* 'dwell, stay', etc.; for *\*e*, cf. Ur. *\*esV-* id.

ii. Nostr. *\*ʔitḁ̄* 'eat' > IE *\*Hed-* id. (Nostr. *\*t* > IE *\*d*), cf. Afro-Asiatic *\*ʔjt* 'eat', Alt. *idḁ̄-* id., as posited

by Illič-Svityč.

iii. Nostr. \**ḡandV* 'man' > IE \**Handh-* id. (cf. reflexes in Gk, Hitt.), cf. Afro-Asiatic: Cushitic \**ḡVnd-* 'people, family, kin'; note Nostr. \**d* > IE \**dh*.

iv. Nostr. \**ʔanSV* 'favorable, friendly' > IE \**Hans-* 'favorable' (as in Gmc. \**ans-*, Hitt. *ass-u* 'good', etc.), cf. Afro-Asiatic \**ʔnS* 'friendly'.

v. Nostr. \**ha* demonstrative pron. > IE \**Ha* id. (Hitt. *a-*), Kartv. \**ha*, cf. Afro-Asiatic \**h-*, Drav. \**ā/\*a-* id., etc. Illič-Svityč reconstructs Nostr. \**ʔa*, though his notes, including his phonetic tables, indicate \**ha*.

vi. Nostr. \**horV* 'arise' > IE \**Hor-* id. (as in Hitt. *ar-*, etc.). Hitt. *a* can originate from IE \**a* or \**o*, but in this case \**o* is supported by data from other IE dialects. Alt. \**orV-* 'arise, enter', as posited by Illič-Svityč and others.

Examples with an IE "strong" laryngeal < Nostr. \**q*, \**ḡ*, \**x*, \**ḡ* :

i. Nostr. \**qantV* 'front' > IE \**Xant-* id. (Hitt. *hant-* 'front part'); Nostr. \**q* is reflected by Afro-Asiatic \**xan* [ḡ] 'front', also 'southern part, nose'; Nostr. \**a* is supported by Alt. \**antV* 'front part, southern part', cf. Illič-Svityč in Ètimolḡija 1965 (1967), p. 354.

ii. Nostr. \**ḡUcHA* 'bone' (\**U* = \**o* or \**u*) > IE IE \**Xost(H)-* id., as in Hitt. *hasta-*, Arm. *oskr*, but it appears that in Armenian, *h-* reflected a voiceless strong laryngeal, and lack of *h-* reflected both voiced strong laryngeal and "weak" \**H*,<sup>4</sup> cf. Afro-Asiatic \**ḡḡ* id., according to Dolgopolsky, and both Indo-European and Afro-Asiatic data reflect Nostr. \**o* rather than \**u*.

iii. Nostr. \**ḡulV* 'wage, war, attack, kill' > IE \**Xwel-*, as in Hitt. *hulla/i-* '(nieder)schlagen', cf. Nostr. \**u* > Afro-Asiatic \**w*, cf. Arabic *ḡwl* 'attack, kill', Cush. \**ḡwVl(l)-* 'wage, he at war', according to Dolgopolsky.

There are only a few examples of IE \*X before a front vowel, e.g. Hitt. *ḫekur* 'summit', *ḫe/ink-* 'perish', etc., which Dolgopolsky has explained as follows: strong laryngeals or (post)velar fricatives tend to appear before [a] or [u], but not before [e] or [i].

Now, let us discuss, however briefly, a recent work by Bomhard (1984), who, while oblivious of the background argumentation of and most of the literature on Nostratic, not to mention an apparently inadequate grounding in Indo-European studies, attempted to demonstrate a genetic relationship between Indo-European and Afro-Asiatic. (As an aside--and hardly to be construed as a gratuitous swipe--we are rather amazed that Bomhard's study was passed for publication: the review process, if there ever was one, could hardly have been very stringent.)

Where IE stems containing voiced aspirates (type D; after Bomhard, IE \**d*, \**g*, etc.) are compared with Afro-Asiatic counterparts containing \**b*, \**d*, \**g*, his results are plausible, and the semantic correspondences seem well founded. But it is precisely here that Bomhard's identification of Indo-European and Afro-Asiatic corresponding phonemes coincides exactly with that posited by both Illič-Svityč and Dolgopolsky, that is, IE \**dh* : AA \**d* Nostr. \**d*, and so on. Naturally enough, then, most of Bomhard's cognates merely repeat those proposed by both Dolgopolsky and Illič-Svityč some two decades ago, and these are, it might be pointed out, also cognates that have been confirmed by congeners in several other Nostratic languages. The principal methodological objection to Bomhard's work is that it is a strictly binary comparison, and, as pointed out above, the margin of error in such comparisons is appreciably greater than in comparisons based on a large number of (related) languages. Moreover, from the Nostratic point of view, a binary comparison of Indo-European and Afro-Asiatic alone is particularly open to misinterpretation, for, as one of the basic postulates of Nostratic theory claims, the original system of Nostratic vowels was greatly confused in both Afro-Asiatic and Indo-European.

In other cases, however, and they are the vast majority, Bomhard's findings contradict those of both Illič-Svityč and Dolgopolsky, and his comparisons here are, by whatever yardstick one cares to use, mostly wrong, and this is easily

seen by the lack of plausible semantic matchings. For instance, Bomhard (1984:207 #50) reconstructs AA *\*t̪ k'-/ \*tak'-*, where *\*tq = \*t̪ k'*, as having the gloss 'touch, push, strike', although the facts from Afro-Asiatic deny such a gloss, cf. Egyptian *tq* 'cut, slay, strike', which, to be fair, is information even Bomhard provides. For Bomhard, then, even self-contradiction is a virtue. However, the value 'touch' in Afro-Asiatic, though it is absent here, is crucial for Bomhard. It gives him the opportunity to compare this Afro-Asiatic radical with IE *\*tag-* (which Bomhard rewrites as *\*tak'-*). IE *\*tag-* meant only 'touch, stroke', meanings lacking in Afro-Asiatic. To make the IE radical "more comparable," Bomhard describes it as having the values 'touch, strike, push, stroke', and he even goes so far as to equip Lat. *tangō* with the values 'touch, strike, push, hit', although Pokorny's (1959) dictionary, apparently Bomhard's only source of information, defines *tango* as *berühren* only. Virtually all of Bomhard's other examples smack of this same hasty, stick and paste, cavalier attitude toward the well-established and well-known comparative and etymological facts of life. Take, for example, his AA *\*t' ɔ y-/ \*t'ay-*, which he (1984:211-2, #65) glosses as 'be bright, pleasant, agreeable', though the meaning 'bright' is totally absent for this root; and this is even clear from the data which Bomhard himself provides. The source of the error is obvious. An IE "cognate" *\*t' ɔ i-/ \*t'ai-*, that is, a root beginning with *\*d-* in the traditional transcription, indeed means 'bright, shine', and Bomhard has missapplied the semantic data for his Afro-Asiatic root.

From these examples--and there are many more like them--we see that Bomhard bases his comparisons on the new (Gamkrelidze-Ivanov-Hopper) interpretation of Indo-European stops. Thus it is that Bomhard compares IE *\*d*, *\*g* in the traditional transcription, his *\*t'*, *\*k'* respectively, with glottalized stops in Afro-Asiatic. However, according to both Dolgopolsky and Illić-Svityč, IE *\*d*, *\*g* derive from Nostr. *\*t*, *\*k* respectively. Hence, one should compare AA *\*t*, *\*k̥*. As for AA *\*t'*, *\*k'*, they putatively derive from Nostr. *\*t*, *\*k'/\*q'* and are therefore comparable with IE *\*t* and *\*k̥*, *\*k*, *\*kʷ*; Bomhard should have compared IE *\*t*, *\*k̥*, etc., with AA *\*t'*, *\*k'/\*q'* and not, as he does, with AA *\*t*, *\*k*.

Quite indirectly, however, Bomhard's blunders both bolster Nostratic theory and reveal that the new (Gamkrelidze-Ivanov-Hopper) interpretation of Indo-European voiced stops is in need of remedial correction. But there are numerous other errors in Bomhard's study as well: a methodologically incorrect reconstruction of Afro-Asiatic and other phonemes on the basis of IE only and reconstruction of IE laryngeals and other phonemes on the basis of AA only; unfounded "chopping off" of parts of AA or IE roots, categorizing is one quite different set of roots of a given language, etc. Presumably, we will have the opportunity to express our reservations and objections in greater detail in the volum entitled, *Die Laryngaltheorie und die Rekonstruktion des indogermanischen Laut- und Formensystems*, currently planned by Alfred Bammesberger.

What follows is a select list of comparisons Illic-Svityc made some twenty years prior to Bomhard's. These examples show that initial IE *\*k*, *\*g*, *\*gh* (in the traditional transcription) remain where we find a corresponding Nostratic or Kartvelian *\*k'/\*q'*, *\*k*, *\*g* before *\*a*, but in Afro-Asiatic we here find *\*k'*, *\*k*, *\*g*, and Nostratic *a*-vocalism is reflected by East Nostratic languages:

1. (i). IE diminutive *\*-k* : Kart. *\*-k'*- id., Ur. *\*-kka* id., where Nostr. *\*k'*, and not *\*q'*, is shown by Kartvelian.
- (ii). IE *\*gem-* 'grab, sneeze' : AA *\*km-* id. < Mong., Tung. *\*kamu-* id.; Nostr. *\*kamu* id., where both vowels have been reconstructed in line with the evidence from Altaic.
- (iii). IE *\*gher-* 'thorn, thorny branch' : Tung. *\*gara*; Ur. *\*kara* 'thorn, branch, conifer'; Drav. *\*kara-* 'thorn, sharp point' < Nostr. *\*gara* 'thorn, thorny branch'. Tungusian is considered archaic as far as Nostratic gutturals are concerned: Nostr. *\*k'/\*q'* > Tung. *\*x*, though probably via an intermediary stage with *\*kh*; Nostr. *\*k* > Tung. *k*; Nostr. *\*g* > Tung. *\*g*. In this case, neither Kartvelian nor Afro-Asiatic is required to indicate Nostr. *\*g*, but note that Nostratic initial gutturals merged in Uralic and Dravidian.

The following are examples of instances in which the Indo-European palatals,  $*\hat{k}$ ,  $*\hat{g}$ ,  $*\hat{g}h$ , correspond to the same stops in Nostratic languages as in the above cases, but where, unlike the above cases, the following vowel was a palatal in Nostratic, i.e.,  $*e$ ,  $*\hat{a}$ , or  $*i$ . (Hence, palatalization in Indo-European is thought to have taken place after all Nostratic vowels not preceded by laryngeals (as per the above remarks) merged as  $*(i)e$  Proto-Indo-European. .

2. (i). IE  $*\hat{k}es-$  'cut': AA  $*qs$  (=  $*k's$ ) 'cut, break'; Turkic  $*k's\hat{s}\hat{s}\hat{a}$  'cut'; Uralic  $*k\hat{a}\hat{c}V$  or  $*k\hat{e}c\hat{a}$  'knife' < Nostr.  $*k'\hat{a}\hat{c}\hat{a}$  'cut'. However, it should be pointed out that reconstruction as  $*k'$  would be impossible without Kart.  $*k'\hat{c}$  (=  $*\hat{k}\hat{c}$ ) 'cut', as only Kartvelian distinguishes between  $*q'$  and  $*k'$ .

(ii). IE  $*\hat{g}l\hat{o}u-$  'brother's wife': AA  $*kl(l)$  'son's wife'; Ur.  $*k\hat{a}l\hat{u}$  'brother's wife' < Nostr.  $*k\hat{a}lU$  'female relative by marriage' where  $*U = *u$  or  $*\hat{u}$ .

(iii). IE  $*\hat{g}herH-$  'dawn, shine': AA  $*ghr$  'day, daylight'; Mong.  $*g\hat{e}re$  'dawn, morning light' < Nostr.  $*gEhra$  or  $*gErha$ , 'dawn'.

Finally, we present examples with consonants from the third triad, namely, IE  $*k^w$ ,  $*g^w$ ,  $*g^wh$ , the labiovelars. They are thought to have originated from the above Nostratic consonants in words in which these consonants preceded a labial vowel in Nostratic, particularly in East Nostratic languages:

3. (i). IE  $*k^w\hat{e}i-$ ,  $*k^w\hat{e}iH-$  'to rest': AA  $*qwj$  (=  $*k'wj$ ) 'to remain'; Ur.  $*koja-$  'lie, rest' < Nostr.  $*k'oja$  or  $*q'oja$  'to rest', where the absence of a Kartvelian congener does not permit us to determine whether we here have underlying  $*k'$  or  $*q'$ .

(ii). IE  $*g^w\hat{e}n-$  'wife, woman': AA  $*kwn$  'one of the wives, woman': Turkic  $*k\hat{u}ni$  'one of the wives' < Nostr.  $*k\hat{u}ni$  'wife, woman'.

(iii). IE  $*g^w\hat{e}r-$  'flame, blaze, burn': AA  $*gwr$  'fire, coals'; Tung.  $*gur-gi-$  'flame, blaze, burn' < Nostr.  $*gUrV$  'live coals', where  $*U = *u$  or  $*\hat{u}$  and the value of the second vowel in the radical is unclear.

These examples are, of course, merely representative and could be multiplied in legion. In 1(i-iii) we have the Indo-European results of Nostr. \**k'/q'*, *k*, *g* before Nostr. \**a*; in 2(i-iii) the Indo-European results of the same series before Nostratic front/palatal vowels; and, finally, in 3(i-iii) the Indo-European results of the same Nostratic series before what was labial vowel in Nostratic. These examples are, for the most part, taken from Dybo's excellent article in *Izvestija*; and Dybo, in turn, took them from the first volume of Illič-Svityč's etymological dictionary of Nostratic, a compact encyclopedia of our current knowledge of the history of the Eurasian languages, languages which are, as Illič-Svityč so brilliantly showed, distantly related children of the same macro-family.

### **Kartvelian**

Before Illič-Svityč began reconstructing Proto-Nostratic, he made several significant contributions to the reconstruction of the Indo-European, Altaic, and Kartvelian proto-languages; and all of these contributions were highly praised by the specialists in these many areas of expertise. Here, however briefly, we would like to examine a paper by B. A. Počxua, "Georgian Lexemes in the 'Nostratic Dictionary'" (*Voprosy jazykoznanija* 6 (1974), 100-105), which is directed against Nostratic theory, as least as far as the inclusion of Kartvelian in Nostratic is concerned, and which was written by a Georgian linguist. This little paper seems to have escaped the notice of most scholars, including ours, until our colleague, Alexis Manaster-Ramer, called our attention to it. (Here we would like to thank Alexis Manaster-Ramer, not only for this reference, but also for many fruitful discussions and much encouragement during our seemingly unending labors on this volume.)

To summarize in brief, Počxua makes a few critical remarks about the use of Georgian material in Illič-Svityč's works. It is clear that he regards Kartvelian languages as being more closely related to North Caucasian languages than to Indo-European, Afro-Asiatic or the other languages subsumed by Nostratic as a macro-family. He reproaches Illič-Svityč for not using North Caucasian data and contends that these languages "are undoubtedly closer to Kartvelian languages than the latter are to Dravidian, Uralic, or Altaic." This assessment must now be rejected in light of recent findings. As Starostin and Nikolaev, as well

as others, have shown by their reconstructions, North Caucasian languages, together with Sino-Tibetan, Yenisseyan, and Na-Dene languages (as well as, perhaps, some others) may all be considered as belonging to a Dene-Caucasian macro-family, whereas Kartvelian distinctly belongs to Nostratic. As for the many lexical correspondences between North Caucasian and Kartvelian, the majority of them are clearly loans, though a few sets might be explained as indices of a very faint relationship between Dene-Caucasian (to which North Caucasian belongs as a daughter language) and Nostratic (of which Kartvelian is considered a daughter language). This is, in fact, the view expressed by Dybo in his (1984:5) introduction to the latest fascicle of Illič-Svityč's dictionary. One might find similar indications of a very remote relationship between, say, Penutian and Uralic; but, despite the efforts of Sadowszki, Webb, and Ivanov toward such a goal, such indicators in no way imply that Uralic and Penutian were ever closely related. Instead, such indicators would, at best, show that Amerind (of which Penutian is a daughter language) was very distantly affiliated with Nostratic (of which Uralic is a daughter language).

Pocxua (1974:102) points out that the several instances where Illič-Svityč compares a Georgian stem with a corresponding stem in but one Nostratic language (e.g., Indo-European, or Uralic, or Afro-Asiatic) are inconclusive. To this objection, one could point out that Illič-Svityč was not comparing just Georgian with Indo-European, or Uralic, etc., but that his basis for comparison was a reconstructed Kartvelian (of which Georgian is but one daughter language). Next, one should note that, if both the phonetic and semantic correspondences are correct (as, indeed, they are), then the correspondences themselves are also acceptable, provided, of course, one is not dealing with cultural loans. Certainly, comparisons between many languages yield more reliable results than do comparisons based on only a handful of languages, or even only two languages, and the vast majority of comparisons in Illič-Svityč's work are based on many, rather than few, languages.

Pocxua also objects to Illič-Svityč's comparison of IE *\*mēg(h)-* (so Pocxua, Illič-Svityč gives *\*meġ(h)-* 'big' with Drav. *\*ma* (so Pocxua, Illič-Svityč gives *\*mā*) 'great'

and Kartv. \**mag-ar* (so Počxua, Illič-Svityč gives \**mag-ar-*) 'big' (Počxua contends that the meaning might well have been 'strong', as in Georgian *magar-*). He says: "In all three forms, only *ma/me* is a common element, and the meanings are rather far apart." Ordinarily, across a wide variety of both related and unrelated languages (literally hundreds of examples could be cited), 'big' and 'strong' go together, and 'big' and 'great' are virtually identical semantically. As for the common element *ma/me*, Počxua seemingly did not bother to consult the phonetic tables in Illič-Svityč's works: they show that IE \**gh* corresponds to Kartv. \**g* and, intervocalically in Dravidian, to zero-consonant with compensatory lengthening of a preceding vowel. In all three cases, this vowel derives from Nostr. \**a*, which regularly shifted to \**e* in Indo-European. The Nostratic proto-form is \**magV* 'big', which Počxua cites, but without any indication of its semantic value. This is highly selective criticism; it is both unfair and counterproductive.

### Phonetic Correspondences

If we compare Nostratic reconstructions from the 1960's with those from recent years, we will find many differences. Dolgopolsky has, for example, made the reconstruction of Nostratic laryngeals and pharyngeals much more explicit: where Illič-Svityč had only \**h* (as in Afro-Asiatic), Dolgopolsky reconstructs both \**x* (as in Kartvelian) and \**h* (Nostr. \**x*, \**h* > AA \**h*); Illič-Svityč's \**q* (earlier \**γ*) has now been replaced by \**γ*, \**q* (both merged in Afro-Asiatic as \**q*). Dolgopolsky also revised some notational conventions: his \**š* corresponds to Illič-Svityč's \**s*, and vice versa (also his \**č* to earlier \**c*, \**ĉ*, to \**c'*, \**ĝ* to \**3*). Dolgopolsky now reconstructs the systems of sibilants and fricatives somewhat differently for both Afro-Asiatic and Altaic; and the latter is, for him at least, not a single, but three different proto-languages (Tungusian, Mongolian, and Turkic). Of note here is the difference between the European and American notation of glottal stops: \**k̥* in European studies corresponds to \**k'* in American studies. In Afro-Asiatic studies, *x* is now generally used for *h*, *γ* for *g* or *g*, and *k̥* (American *k'*) for *q*. Note that Hittite-Luwian \**x*[*x*] is transcribed as \**h* or \**h*.

Then, too, there are differences in the reconstructions

proposed for the Nostratic daughter languages. Diakonoff reconstructs labio-velars for Afro-Asiatic, while Illic-Svityc and Dolgopolsky do not. Some changes have to be made in what is assumed to have been underlyingly there in the proto-language as the result of recent discoveries about the Cushitic and Chadic languages; this is also true of Uralic and the Caucasian languages.

We may now summarize our proposed changes in the reconstruction of the Indo-European obstruent system as follows:

### **Classical Nostratic**

Nostr. laryngeal, pharyngeals,  $*ʔ$ ,  $*q$ ,  $q > IE *h$ , which became  $*h̥$  before an original front vowel, remained as  $*h$  before  $*a$ , and became  $*h^w$  before an original labial vowel.

### **Our Proposal**

The Indo-European laryngeal as not affected by a neighboring vowel, and the original quality of the following vowel was preserved, thanks to the preceding laryngeal. Nostr.  $*e$ ,  $*ḁ > IE *e$ , Nostr.  $*i > IE *e$  or an  $i$ -diphthong; Nostr.  $*a > IE *a$ ; Nostr.  $*o$ ,  $*u$ ,  $*u̥$ , Illic-Svityc reconstructs both Nostr.  $*u$  and  $*u̥$ ; Dolgopolsky now reconstructs only  $*u$ ,  $> IE *o$  or a  $w$ -diphthong; Nostr.  $*q$ ,  $*g$ ,  $*x$ ,  $*γ > IE$  strong  $*X$  probably with the following differentiation:  $*q$ ,  $*x > *x$ , but  $*g$ ,  $*γ > *γ$ ; Nostr.  $*ʔ$ ,  $*q̥$ ,  $*h$ ,  $*h > IE$  weak  $*H$  (probably more than one variety of same).

### **Classical Nostratic**

Nostr.  $*T'$ ,  $*T$ ,  $*D > IE *T$ ,  $*D$ ,  $*Dh$  respectively.

**Our Proposal** (a modification of that by Gamkrelidze-Ivanov-Hopper) – see *supra*, p. 15.

It seems possible that IE  $*T$  was  $*Th$ ,  $*D$  was  $*T$ , and  $*Dh$  was  $*D$ , as in Altaic, so that we have Nostr.  $*T'$ ,  $*T$ ,  $*D > IE *Th$ ,  $*T$ ,  $*D$  respectively (as in Armenian) with a "strong"  $Th$  and a very breathy  $*D$ . Certainly, the second IE stop was not glottalized.

Finally, then, it is important to underscore the fact that Nostratic reconstructions provide a control mechanism for strictly internal comparisons that would otherwise lack such control and, perhaps most important, Nostratic would seem to yield explanations for highly archaic

linguistic phenomena that would otherwise lack satisfactory explanations.

An example of the latter situation is provided by IE *\*a<sub>1</sub>es-* 'mouth', as in Hitt. *aiš* (< IE *\*a<sub>1</sub>es*, so Pedersen), and *\*ōus-* id. (as traditionally transcribed) were correctly combined as cognates by many comparatists, but they were then troubled by the lack of a satisfactory explanation for this combination and subsequently began to offer implausible explanations and view them as separate etyma (e.g., Hitt. *aiš/išās* 'mouth' < IE *\*H<sub>1</sub>es-os*; Hitt. *aiš* related to *has(š)-/heš(š)-* 'to open'). A plausible explanation is supplied by Nostratic theory, according to which the Nostratic radical was *\*?iw3V* 'to swallow' (so Dolgopolsky) with regular variation from Nostratic radicals with *\*-iw-* such that *\*Hius-* > *\*Hāiws* > *Hōus* (with compensatory lengthening upon elimination of *i*) vs. *\*Haiws-* > *\*Hais-*, etc. (with disappearance of *\*w*, as Indo-European did not tolerate structures of the type *\*-χws-*). Here, note that the Indo-European "weak" laryngeal (from Nostr. *\*?*) regularly disappeared in Hittite and that IE *\*o*, but not *\*e* (from Nostr. *\*i*) often appeared in roots that had *\*w* in Nostratic and in Indo-European etyma categorized as nouns.

This is but one example of how Nostratic theory can push the cloud of developmental obscurity back in time. But even for comparatists not conversant with Nostratic theory, such as Philip H. Baldi, Indo-European has taken on a new look and assumed a vastly different appearance from what it had even some thirty years ago, and the canonical shapes postulated by Karl Brugmann (1849-1919) and his generation have necessarily been greatly revised, but then, so too, that generation of neogrammarians, particularly Saussure (1857-1913) among them with his sweeping revision of the vowel system, compellingly revised the knowledge it had inherited from Franz Bopp (1791-1867) and Jacob Grimm (1785-1863).

Thus it is that our understanding of Indo-European, not to mention Nostratic, and its increasingly earlier forms and ever more distant affinities, has been a gradually cumulative, rather than an abruptly cataclysmic, affair. But it is not only in the realms of phonological correspondences and the structure of the root that our knowledge has broadened in space and deepened in time. From recent investigations of the minor, so-called "*Trümmer-sprachen*" (e.g. Phrygian, Lycian, Carian, Osco-Umbrian) we

now realize that the arsenal of grammatical strategies at the disposal of the *Grundsprache* must have been far richer than classical Indo-European theory had ever assumed. We still--and probably always will--lack a complete inventory of the Indo-European (or, for that matter, Nostratic) lexicon; but here, too, we have benefited enormously from the investigation of *Trümmersprachen*. As our knowledge of Indo-European expands, so too will our knowledge of Nostratic, particularly in its grammatical dimension. Many, many problems remain, however; and they are problems that are dramatically different, in magnitude, if not in kind, from those our forebears confronted. Perhaps there are phonological and morphological regularities that are still unknown to us, or even underlying morphological developments not yet fully appreciated or understood in a typological and developmental sense, that were obliterated in part, or, conversely, that have obliterated what were originally exponentially more numerous points of similarity. We now realize, for example, that ergative languages are heavily out-numbered by non-ergative languages; and it seems to be the case that, in contact situations, ergativity tends to give way to non-ergativity, but the reasons for this are as yet largely unknown. Then, too, it is very likely that the unstable portions of the lexicon were largely renewed after periods of common development, and, given the core of stable items identified by Dolgopolsky, it remains for us to discover why it is that precisely these, and no other, items are universally stable. We may, in fact, be addressing problems that lie at the very heart of human linguistic cognition. If so, then the search for Nostratic and other proto-languages (Dene, Caucasian, Amerind, Macro-Asiatic) will take on a new urgency and a new significance.

## NOTES

### Shevoroshkin - Foreward

1. Here are some examples which show some of the differences between American and Soviet linguists with

respect to broader genetic relationships between languages. American linguists have debated whether or not there is such a thing as a genetic unity for "Hokan languages", and most of them have asserted that there is not, while Soviet linguists have proceeded to reconstruct Proto-Hokan. Accordingly, one might expect Soviet linguists to be busily engaged in reconstructing Proto-Penutian while their American colleagues, with but few exceptions such as Victor Golla and Michael Silverstein, are still stalwartly denying the existence of such a genetic unit, even though there appears to be a well-defined Penutian family of languages demarcated by phonetic correspondences in both grammatical elements and lexical items, a family that includes Miwok-Costanoan, Maidu-Misenan, Yokuts, Wintu-Patwin, Klamath-Sahaptin, Takelma-Kalapuyan, Chinoon, and Tsimshian, but not Coos, Siuslaw, or Alsea. Nikolaev and Leščiner have asserted that the Proto-Penutian phonological system is virtually identical with that of Proto-Hokan. Moreover, both Proto-Penutian and Proto-Hokan have identical sets of the stablest morphemes (e.g. personal pronouns), and may be remotely related as members of the "Amerind" macro-family as this has been defined by Josepy H. Greenberg. We note, however, that, in his 1980 report to the Summer Meeting of the Linguistic Society of America in Albuquerque, Greenberg incorrectly included some North American languages in "Amerind", viz. Algonquian-Wiyot-Yurok, Yuchi-Siouan, Caddoan, Iroquoian, Salishan, Wakashan, all of which belong to an altogether different macro-family. Greenberg correctly defined Amerind languages as those languages with  $nV$  and  $mV$  for the 1st and 2nd person personal pronouns respectively. Both O. Sadowszki and V. Ivanov have erroneously defined Penutian as closely related to Uralic, which, in turn, belongs to the Nostratic macro-family, cf. Shevoroshkin (1981).

2. In Dolgopolsky's later paper, "Soxran'aemost' leksiki, universalii i areal'naja tipologija" (Degrees of Lexical Stability, Universals, and Areal Typology) in *Lingvisticeskaja tipologija i vostochnye iazyki* (Moscow: Nauka, 1965), pp. 189 ff., he comments on some rather insignificant exceptions to the items on this list, particularly the pronouns, 'I/me' and 'thou/thee', and shifts such as English *you* instead of *thou*, as well as other matters that have

to do with the T/V-split. There are a few instances of somewhat lesser stability for some items, e.g. for 'blood' in Indo-Aryan and Dravidian, the meaning of 'ear' in Celtic, and so on. Nevertheless, even as it stands, Dolgopolsky's article in this collection represents a rather dramatic step forward in the direction of codifying a cogent program for comparative studies.

3. Cf. now Dybo's (1984:10) remarks in his introduction to the latest fascicle of Illič-Svityč's dictionary: "The Nostratic hypothesis cannot pretend, as Serebrennikov would have it, that the Nostratic proto-language existed some forty thousand years ago. It is quite certain that Nostratic was not the initial language of *Homo sapiens*, and if objections are made on the basis of this assumption, then they should not be aimed at those who engage in Nostratic research."

4. If we combine the Anatolian (Hittite) and Armenian evidence as Edgar C. Polome has so aptly done in his paper, "Armenian and Proto-Indo-European Laryngeals" (In: *First International Conference of Armenian Linguistics: Proceedings* (Caravan Books: Delmar, NY, 1980), pp. 17 ff.) with the data as we now know it from Nostratic, then we might conclude that both Armenian and Hittite preserved IE voiceless \*x (< Nostr. \*9, \*γ); weak IE \*H was lost in both Hittite-Luwian and Armenian. Indeed, in addition to Arm. *oskr*, Hitt. *haštai-* 'bone' (< IE \*γost(H)- < Nostr. \*γ[o]c'HV 'bone'; IE \*st < Nostr. \*č' and \*č), we find: Arm. *orb* 'orphan', Hitt. *harp-* 'separate' < IE \*γorbh- < Nostr. \*γorbV 'be deprived of (something)' (according to Dolgopolsky); Arm. *ost* 'branch', Hitt. *hasduir* 'weeds, refuse' < IE \*γoz-d- < \*γos-d- < . . . γ < Nostr. \*γoč'V 'branch' (IE \*s < Nostr. \*s, \*š, \*š, \*c, \*č, \*c', \*č'); Arm. *oror* 'gull', *ori* 'rook' (all from 'bird' ?) : Hitt. *hara-n-* 'eagle' < IE \*γVr- < Nostr. \*γVrV, etc.

If so, then we might have Nostr. \*q or \*x > IE \*x > Hittite-Luwian \*h, Arm. \*h as in Hitt. *hanna-*, Arm. *han-* 'grandmother'; Hitt. *huhha-*, Arm. *haw* 'grandfather', Hittite-Luwian \*hawi- 'sheep' : Arm. *hovi-w* 'shepherd', etc.

Weak IE \*H (lost in Hittite-Luwian, Armenian, and other Indo-Languages): Hittite *es-mi* vs. Arm. *e-m* 'I am' (< IE \*Hes- < Nostr. \*?esA, see Illič-Svityč (1971:268-9); Hittite *et-mi* vs. Arm. *ut-em* 'I eat' (with u- < \*ō < \*ē) < IE

*\*Hed-* < Nostr. *\*ʔitā*<sup>h</sup>, see Illič-Svityč (1971:273-4), cf., also, albeit without Hittite-Luwian parallels, Arm. *aviun* 'desire', Lat. *aveō* 'I crave', etc. < IE *\*Haw-* < Nostr. *\*hawV* 'to crave', cf. Illič-Svityč (1971:241-2) and note Afro-Asiatic *\*hw*, *\*hwj* 'to crave', Drav. *\*āv-* id.

If this line of reasoning is correct, then we can reconstruct Indo-European laryngeals even more precisely than was done above:

Nostr. *\*ʔ*, *\*C*<sub>1</sub>, *\*h*<sub>1</sub>, *\*h* > IE *\*H* > Hittite-Luwian, Arm., etc.  $\emptyset$

Nostr. *\*q*, *\*x* > IE *\*x* > Hittite-Luwian *\*h*, Arm. *h*

Nostr. *\*g*, *\*γ* > IE *\*γ* > Hittite-Luwian *\*h*, Arm.  $\emptyset$

We have reconstructed three laryngeals for Indo-European whose quality is independent of that of neighboring vowels and we have done so on the basis of Indo-European data alone: the weak laryngeal disappears in Indo-European, though it regularly lengthens non-initial vowels or consonants in Hittite-Luwian and other languages; strong voiceless "laryngeal" yields Hittite-Luwian *\*h*<sub>1</sub> and Arm. *h*<sub>1</sub>, while strong voiced "laryngeal" turns into *\*h* in Hittite-Luwian, but disappears in Armenian. Having done that, we then compared the results with the Nostratic data and found a basis for confirmation there. The Nostratic data might suggest that there was more than one weak laryngeal in Indo-European, but the data from Indo-European alone so far provide a basis for the reconstruction of but one weak laryngeal, namely *\*H*, for Indo-European.

**PROTO-LANGUAGES**  
**AS OBJECTS OF SCIENTIFIC DESCRIPTION**  
**V. V. Ivanov**

**The Difference Between a Proto-Language and a Mere System of Correspondence**

It is appropriate in both comparative-historical linguistics and other sufficiently developed scientific fields to distinguish between a formal apparatus and its material interpretation. The formal apparatus of comparative-historical linguistics includes, among other things, a technique for establishing a system of phonetic correspondences between genetically related morphs and the reconstruction of the proto-forms of those morphs—their phonetic constituents. A formal system of phonemes and morphs reconstructed in this fashion requires interpretation. Theoretically, a proto-language represents only one such interpretation. Before we label the reconstructed system of correspondences a proto-language, we have to exclude other possible interpretations, which we are going to list below.

First, correspondences may appear as the result of a large number of borrowings, either from one language to another, or from a third language to both languages being compared. Both possibilities can be easily identified when later stages of a language's history are being investigated, provided, of course, some early written records are present (for instance, when mutual morphs of Chinese, Japanese, and Vietnamese are compared). On the other hand, both possibilities can present some real difficulties when earlier stages of a language's history are investigated. (Consider, for example, the fact that, for a long time, the main criterion for positing a genetic relationship between Vietnamese and Thai on the one hand and neighboring languages on the other hand was a stratum of archaic borrowings which were sometimes invoked as proof of genetic relationship.) It is a testimony to the profound insight and outstanding sense of scholarly propriety on the part of V. M. Illič-Svityč that he was able to identify a set of loans from Proto-Semitic in Proto-Indo-European which enabled him to extrapolate this set, as well as morphs borrowed by Indo-European and Proto-Semitic from some one source (particularly Sumerian) from the list of inherited

common Nostratic morphs.<sup>1</sup> In this respect, the Indo-European - Semitic comparisons in the work of Illič-Svityč differ advantageously not only from those in the works of his predecessors, Möller and Cuny, but also from those in the interesting experiments recently published by Bomhard.<sup>2</sup> Bomhard simply failed to take into account the totality of insights accomplished by Illič-Svityč. A case such as that being reviewed here is valid not only if unrelated languages are compared (such as Japanese, Vietnamese, and Chinese), but also for languages, albeit genetically related, which have themselves become widely differentiated, as, for example, the Indo-European and Semitic proto-languages, which, in terms of Nostratic theory, are putatively genetically related. Similar problems can arise when, as the result of contact among closely related languages, a large number of morphs with a high degree of functional load have been borrowed. Well known examples of this are provided by the numerous Iranian loans in Common Slavic. Though there is an indisputable set of regular correspondences between Common Slavic and Iranian, no one would doubt but what most of the reciprocal morphs have to be considered, both formally and functionally, as relatively late introductions into Iranian (therefore, a hypothesis in favor of a Common Slavic-Iranian Proto-Language is completely excluded). As a typical example, one could cite OPers. *rādiy* (>Mod. Pers. *ra* [r : ]), which corresponds exactly, both formally and syntactically, to Sl. *\*radi* in such constructions as Russ. *otčā radi* 'for father's sake', *boḡa radi* 'for God's sake!'. Nevertheless, in old Persian itself this construction has a later areal existence (in the sense of a speech community vs. a language community), though it is just as old from the point of view of absolute chronology. This can be shown by the typological coincidence of the old Persian postpositional *rādiy* with the Elamite postpositional *ikku/i* (e.g., in constructions such as <sup>d</sup>*In-su-si-na-ak na-pir- u-ri ik-k<sup>u</sup>/i tu-um-pa-ah* 'I have established it for the sake of Insusinak, my God', see Grillot (1978:24), as well as with Hurrian *edi* (e.g., in *at-ta-i-ip-pa e-ti-i-i-ta* 'for my father's sake', *Tusratta Letter*, III 52-3), see Speiser (1941):53-4, 72). Thus, we can assume that a Persian postpositional construction of this type, which was absent in Indo-Iranian, appeared under the influence of those Near Eastern languages with

which the Iranians came into contact before a portion of them had migrated to Eastern Europe where they came into intimate contact with the speakers of Common Slavic. The situation with which scholars are confronted when investigating Slavic-Iranian correspondences reminds one of the contacts between Old English and Norse, a contact situation when not only lexical items, but also personal pronouns, were borrowed from Old Norse (which is only possible when closely related languages or even dialects of a language are in intimate contact).

In cases such as those being reviewed here, that is, cases in which the material interpretation of correspondences between languages reveals a vast number of archaic loans with substantial morpho-syntactic functional loads--the very basis for the correlation of correspondences, one is readily induced to assume a *Sprachbund* interpretation. In the case of a *Sprachbund*, there can be not only the revelation of a large number of mutually shared words (borrowed either from one of the members of the *Sprachbund* or from a source common to all the languages in question, e.g., a substrate or superstrate), but also the realization of a mutually shared set of phonemes (specifically, the abruptives in, for example, Ossetic, Armenian, and Kartvelian, all of which are members of the greater Caucasian *Sprachbund*), as well as a set of grammatical phenomena. In a situation in which not only lexical borrowings can be adduced, even partially, as having originated from some one common source (e.g., substrata, or archaic and/or superstrata), but also phonological and syntactic phenomena, then a comparativist finds himself faced with a substantially complicated task, namely, that of delimiting these *Sprachbund*-phenomena from a proto-linguistic heritage. Such a task is, in principle, soluble, because archaic components of a grammatical system, identified by means of internal reconstruction, cannot be considered as grammatical phenomena typical for a *Sprachbund*. Herein lies the distinction between, on the one hand, the genetic relationship of languages originating from a specific proto-language and, on the other hand, so-called "allogenic relationship" (G.V. Tsereteli's term), as theoretically elaborated as early as the 1920's by E. D. Polivanov and subsequently employed by N. S. Trubetzkoy and V. Pisani in their studies of the Indo-European "problem"

(Pisani considers Indo-European a superstratum, typologically comparable to Sanskrit, as an overlay to more archaic indigenous languages, see Pisani [1959].)

Let us consider two examples (both explained, for reasons of simplicity, on the basis of correspondences inferred between two languages belonging to the Nostratic phylum) of such correspondences between grammatical systems which cannot be explained by late contact between languages or by allogetic relationship resulting from contact. In the course of the internal reconstruction of the oldest verb paradigm, conventionally labelled secondary endings in Indo-European, it becomes possible, on the basis of a comparison of the Hitt. *hi*-conjugation and the mediopassives, as well as the perfect and middle in other Indo-European languages, to establish the following set of the oldest desinences that designate non-volitional action (cf. typological parallels in Amerindian languages), see Hill (1969), McLandon (1978):

#### Singular

1. \*-H-w-
2. \*-(H-)
3. \*-(H-?)

Grammatical correspondences with Nostratic as discovered by Illič Svityč show that some morphological elements that have functioned as suffixes in Proto-Indo-European could recur as prefixes in other Nostratic languages. (This is due to a difference in the underlying syntactic constructions from which these forms originated.) In this case, the above paradigm for the non-volitional suffixes coincides formally with the Common Kartvelian paradigm of subject prefixes as reconstructed independently of Indo-European, see Klimov (1977:258):

#### Singular

1. \*-x-w-
2. \*-x-
3. \*(?)

The realization of these paradigms for common Indo-

European is supported by formal coincidences in the following oppositions: IE non-volitional suffix for the 1st pers. sg. *\*-H-w-*: 1st pers. sg. *\*-m-* ~; Kartv. 1st sg. subj. form *\*-x-w-*: 1st sg. *\*-m-*. There is a fundamental opposition in Kartvelian between the subject affixes (Svan *xw-i-ked* 'I take [that]', *x-i-ked* 'thou take [that]', cf. IE *\*bher-o-H-* 'I take', *\*doH-w-* 'I give/take, am in a state of reciprocity as far as giving is concerned') and object affixes. Comparison with the series of the Indo-European verb is possible if we assume that the affixes in *\*-m-* in both Kartvelian and Indo-European originate from affixed pronominal stems.<sup>3</sup> An operational difference between the subject function of IE *\*-m-* and the object function of Kartv. *\*-m-* is considerably reduced if we accept an originally passive meaning for the corresponding Indo-European paradigm.

To explain why the subject function of verbs inflected with *\*-m-* is not original, one has to point out characteristic performative peculiarities for the verbs in this category, namely, their incidence in impersonal constructions such as Lith. *sniegti* 'it snows', Avest. *snaēžati*, Gk. *νειφει*, Latin *nivet*, OIr. *snigid*. As an illustrative example, one can cite the Hitt. 3rd sg. pres. *istark-zi* from *istark-* 'be ill'. This form usually occurs in impersonal constructions of the type *antuḫšan kuin istarkzi* 'the person whom it ill' (KBo xi 74 Rs III 4); *sallinpat kuinki istarkzi* 'whatever/whichever adult it ill' (KBo xxi 100 I); *mān UKU-an KA x U.SU istarkzi* 'if it ill a man (his) mouth' (partitive apposition with accusativus duplex) (KBo xxi 20 Vs I 12). Lith. *niežti mi* 'it itches/cramps me' (Latv. dial. *kam niēz* 'to whom it itches') belongs to the same semantic field as Hitt. *istark-* (cf. Lith. *sėrgu*, Latv. *sērgu*), and this Lithuanian stem is related to Avest. *naeza*, the name of a disease, Oss. *nīz*, *nez* 'disease' in constructions such as *niz kaemae naei*, *wi nae qærzy* 'to whom there is no disease, he does not groan'. The Iranian correspondence substantiates the assumption that the above impersonal constructions originated from the designation of a force hostile to a person and with which that person was involved against his will. In Baltic, verbs meaning cough, sneeze, and other unpleasant physical reactions were formally members of this grammatical type. The archaism of the Greek perception of sneezing as a manifestation of hostile

forces can be demonstrated by a comparison with other Indo-European traditions and typological parallels, see Onians (1954:103-5, 120, 205, 483, 487). To these parallels, data can now be added from Amerindian languages of the Eastern Pomo type in which verbs with similar meanings (e.g., Epomo *?éčki* 'sneeze') belong to a class of verbs designating action beyond the speaker's control and volition, see McLendon (1978: 4, 6). Such Eastern Pomo verbs are, as a rule, used without agents and are connected solely with the designation of a patient who undergoes physical reactions uncontrolled by his volition.

In light of these typological comparisons, the Common Indo-European use of verbs typified by *\*-m-* such as *\*es-* 'to be' (e.g., Lat. *est mihi filius* lit. 'there is a son to me', 'I have a son') appears to be rather archaic.

A typological rapprochement with the Yenisseyan languages enables us to assume that, in the verbal suffixal series *\*-m-* (*\*-s-*, *\*-t-*), Proto-Indo-European developed personalized verb forms from possessive constructions with pronouns: Ket. *batabdāk* 'my extracting it' → > 'I will extract it', see Kreinovič (1968:24). In Yenisseyan languages, this development was predicated on the assumption that, as first suggested by Castrén, verbal prefixes are merely the fragments of pronouns. A similar assumption for Indo-European follows from a venerable observation of this sort by Bopp, who believed that *-mi* originated from a pronominal form. Comparison with Hittite constructions of the type *šer-šit* 'for him' (< 'his upper part'), an earlier expression of inalienable possession, enables us to accept a structural and semantic parallelism between these constructions and originally syntactic combinations that generated forms of the type *\*es-mi* 'I am' from an underlying possessive construction 'being-my'.

A justification for the typological rapprochement between Indo-European verbs of this type with Yenisseyan constructions is substantiated by the fact that verbs of phonation (e.g., 'say, speak, sing') also belong to such a morpho-syntactic class. In Yenisseyan, impersonal possessive constructions are typical, both for the expression of wishing (Ket. *ab-q-oj* 'my wish' = 'I want') and to denote the semantic contents of phonation (e.g., 'say' vs. 'speak; vs. 'sing'). Exact structural and semantic parallels for this can be found in verbs of phonation in Balto-Slavic

and Anatolian. Cf. Hittite impersonal constructions of the type *nu-za išhamaiškizzi* 'and it sings (for itself)' and Yenisseyan 'a sound of song resounds (for itself)'. Semantic reconstruction, as well as the very reconstruction of two primary series of verb desinences, finds additional support from a comparison with two similar series in an East Nostratic language, namely, Uralic. The bases for an internal reconstruction of the Hungarian verb were previously established by Lewy, who underscored the peculiarities of Hungarian and the old Ugric object conjugation, peculiarities that had been explained as having earlier derived from a possessive, see Lewy (1961:7, 383, 488, 491). E. A. Xelimsky's (1982:86-8) comparison of the Ugric and Samoyedic verb made it possible to reconstruct two series of forms that coincided in principle with both Indo-European and Kartvelian. (Particularly illustrative here is the Hungarian opposition between the 1st pers. sg. in *-k* of the subject conjugation and the 1st pers. sg. in *-m-* of the object conjugation; cf. too, the 3rd pers. *-k-* of the conjugation in *-ik*.) All of this serves to substantiate an earlier suggestion about the possibility of a regular correspondence between *\*-H-* in the above Indo-European paradigm and Uralic *-k-*, see Rosenkranz (1966), but with a very real difference: here, we have to do, not with a common Indo-Uralic phenomenon, but with a common Nostratic grammatical archaism (ultimately syntactic on the Nostratic level). This last conclusion can be validated both by the above striking Kartvelian-Indo-European similarities and the pervasive coincidence of the paradigms of the verb 'know' (sg.): IE *\*w(o)id-H-* (cf. synonymous Hitt. *sakhi* with the same marker *\*-H-*) and Sem. *\*jd?*:

Person	Old Akkadian	Indo-European
1.	<i>īd-e</i>	<i>*woid-H<sub>2</sub>-e</i>
2.	<i>t-īd-e</i>	<i>*woid-t-H<sub>2</sub>-e</i>
3.	<i>īd-e</i>	<i>*woid(-H<sub>2</sub>)-e</i>

The verbal ending in the Indo-European series of the 2nd pers. sg. *\*-tH<sub>2</sub>e* > *\*-tHo* (unstressed, otherwise *-a*) can be explained as consisting in an agglutinative combination of the secondary marker *\*-H-* added to the pronominal

element *\*-t-* of the same origin as the personal and possessive pronoun of the 2nd pers. sg. in *\*-t-*. On the basis of functional correlations, identification of the Old Akkadian (still close enough to its Common Semitic source) paradigm with the Indo-European paradigm can now be proposed. In both languages, the paradigm is unique; here, internal reconstruction enables us to rank it among those archaic features which cannot be explained alternatively, that is, by language contact or "allogenetic relationship." In Semitic (in our case, Old Akkadian), only the preterite used as a stative can be derived from *jdʔ*. It corresponds fully in function to the Indo-European form of the second series *\*voidH<sub>2</sub>e* in those languages (e.g., German *weiss*) where this form is preserved as a preterite-present or the like (cf. Latv. *nevaids* < *\*ne vaids*, OCS *vědě*, etc.). The paradigms can be correlated by virtue of the identity in desinences. The 2nd pers. displays the typical distinction between prefixation of (originally pronominal) *\*t-* in Semitic and suffixation (= infixation before *\*-H-* and after the root) in Indo-European. The absence of *\*w-* in Sem. *\*jdʔ*, in opposition to IE *\*void-H-*, can be compared to peculiarities in the conjugation of the Akkadian neutral statives (i.e., verbids), such as *warāqum* 'to green, be green' (cf., similarly, lack of *w-* in *irriq, īriq, itariq*) and with the formation of the perfect without *w-* (with the initial *t-*, i.e., with another archaic prefixal morph) from verbs of the type *wabūlum* 'carry/bring' (*ittabul* 'he brought', *itbal*, etc.), *wṣī* 'leave/walk out' (*ittazī* 'he left/walked out'). In other words, initial *w-* is simply absent in stative or perfective forms; as for the typology of the morpho-semantics of such verbs as Old Akkadian *warāqum*, one may compare Amerindian verbs of the Cupéño type *k<sup>w</sup>ati-yaxe* 'be red', which always belong to the "non-volitional" (see Hill 1969:355) or "inactive" (cf. Klimov 1977) category of verbs. On the basis of the formal and semantic arguments presented above, one can infer, given acceptance of the traditional explanation of *\*o* as a zero-grade reflex, that the Indo-European and Semitic paradigms are identical. Actually, these facts alone are sufficient to assert that the Nostratic relationship cannot be an "allogenetic" one.

As a second grammatical corroboration of the reality of the interpretation of the correlations between Nostratic languages construed as traces indicating a common origin,

indeed, as evidence of one proto-language, identity in the most archaic types of heteroclitic declensions in both Indo-European and Dravidian can be cited. Every archaic Indo-European language, including Hittite, preserves unquestionable traces of that primordial state, that exceptional archaism, of heterocclisis in neuter (inanimate) nouns in which there was alternation between affixes containing *\*-n-t-* in the oblique cases vs. affixes without this element in the direct cases. Illič-Svityč accepted (cf. Ivanov 1959:24-7) the hypothesis that Tocharian nominal paradigms reflected a common Indo-European protolinguistic structure, see Illič-Svityč (1976-79). He (1976:81) was the first to compare heteroclitic paradigms with the formant *-n-* in Indo-European and Dravidian. According to this hypothesis, "at an early stage in Proto-Dravidian, this formant functioned as a marker of an undifferentiated oblique form" (1976:80). The arresting ingenuity of this insight by the founder of Nostratic historical-comparative grammar as early as the mid-1960's has now been corroborated by the research of the 1970's. It was found that Proto-Dravidian impersonal nouns (which correspond functionally to Indo-European neuters) and those in *-r-* had an increment in *-tt-* < *\*-n-tt-*, see Shanmugan (1971:245). It is possible, as indicated by other Elamite-Dravidian correspondences, that Elamite *-n-*, *-t-* in inanimate nouns also belong here, see Grilhot (1978:3). There can be no doubt but that the Common Dravidian nominal type in (*\*-r-/*) *\*-n-tt-*, preserved in several Dravidian languages as a protolinguistic reflex, corresponds exactly to the Indo-European heteroclitic type in (*\*-r-/*) *\*-n-t-*, also preserved only in isolated relics in some older dialects, e.g., Lat. *iter*, *itineris* 'road'. For striking Dravidian parallels, cf. Kota *ynār* 'road' with *-r-* < *\*-r-*, Tamil *arū* 'way'. As for the verbal counterpart, 'walk/go', Lith. *einū* 'I walk' is important, for its archaic nature is borne out by parallel forms in Latin, Tocharian, and Anatolian. This Indo-European - Dravidian correlation, supported by further parallels in both Kartvelian and other Nostratic languages and uncovered by Illič-Svityč, is particularly convincing since, in both Proto-Indo-European and Proto-Dravidian, heterocclisis is independently reconstructed as a highly archaic grammatical feature which can be projected onto the very core of their grammatical systems by means of internal reconstruction. This fact alone

excludes any presumption that the typologies considered here could have emerged in both Dravidian and Indo-European as the result of an allogenetic relationship.

Finally, it is worth pointing out that both of the examples reviewed here, in which the establishment of correspondences between grammatical archaisms excludes any explanation of systematic correlation based on an hypothesis of allogenetic relationship, are mutually connected by the very morpho-semantics of their corresponding correlations.

As shown by a typological comparison with Amerindian languages of the Eastern Pomo type, an opposition between "volitional" and "non-volitional" verbs (in other words, *active : non-active*) emerges, as a rule, precisely in those language systems where a distinction between nouns of animate and inanimate (i.e., *active : non-active*) classes is typical.<sup>4</sup>

The larger the body of irrefutable grammatical evidence for the primordial relationship among the Nostratic languages becomes, the more substantial the mutuality of shared lexical phenomena becomes; and this evidence points to the preservation of archaisms from a Common Nostratic period. As an illustrative example, one can cite Hitt. *neka*- 'younger female relative, daughter, sister',<sup>5</sup> an item which must have been accepted as archaic for several reasons. First, this term reveals a clear trace of a classification based on age and unites different generations (daughter - sister, both *heterostathmique* and *homostathmique*), an especially significant feature for kinship systems of the Omaha type, and this kinship system is indicated for Indo-European by virtue of several semantic facts. Second, the word for 'sister' in other Indo-European languages may be considered a secondary replacement for a more archaic term;<sup>6</sup> replacement seems to have resulted from a restructuring of the original kinship system. Therefore, this kinship term, an archaism in Hittite that was replaced in Luwian in historical times, can be compared with another kinship term preserved in a different and geographically remote group of Nostratic languages, namely, Tungusian, cf. Manchu *non* (pl. *nota* with the ending *-ta* characteristic for kinship terms), Churchun *niēh-hūn-wēn* 'younger sister', Negidal *nehu* 'younger sister/brother', Evenki *nekū* 'younger sister/brother, younger relative',

a Common Tungusian word which might also correspond to Yakut *noko*, *noho* (form of address by elders to younger relatives), see Tungusian (1975:617-8).

Semantic and phonetic (Luwian shows loss of IE  $*\hat{g}h$ -) correspondences between the word for younger female relative in Anatolian, which retains an archaism from Common Indo-European, and Tungusian (probably also Turkic) seems likely. Borrowing is excluded for geographical and historical reasons.

The discovery of such isolated lexical archaisms for which correspondences are found in remote areas of the East-Nostratic group can, it seems, be expected not only in Anatolian, but also in some other Indo-European dialects whose lexicons have long been considered exceedingly archaic. They have been considered so because of the presence of several Latin/Indo-Iranian, as well as Latin/Anatolian, similarities; that is, similarities that evidently stem from the detritus of a Common Indo-European lexicon, and a lexicon that is mostly of a ritualistic or sacral character. Proceeding in this way, one can, for example, consider the striking similarity between Latin *sanguis* 'blood', so far unexplained, and the Common Tungus-Manchu word reflected in Manchu *senq̄qi* 'blood', Evenki *sēŋi*- 'blood', where, as in the case of Hittite *neka*- : Evenki *nekū* (etc.), borrowing is excluded for geographical and historical reasons.

A qualification has to be made, however, about the presence of phonetically and semantically similar words in two mutually remote groups of Nostratic languages; such incidence cannot, in and of itself, be construed as a full warranty of their mutual Common Nostratic origin. Additional, sufficiently compelling arguments for the archaic nature of the given term must be present (as in the case of the kinship term cited above). Otherwise, an assumption of archaic borrowing may have to be regarded as possibly providing a more plausible interpretation of the correlation. Such is, it seems, the case with the indisputable connection between the Tungusian word for 'mouse, rat' (e.g., Manchu *sinq̄geri* 'rat', Evenki dial. *sin̄er̄ek̄en* 'mouse, rat', Negidal *sin̄ējē* <  $*sin̄er̄ē$  id., Oroch *sin̄e* id. Ulchan, Orok, Nanaj *sin̄q̄ere* id.,<sup>7</sup> on the one hand, and by Old Semitic (as reflected in Akkadian *humuṣṣirum* 'rat') on the other hand. Though the lack of an initial

phoneme corresponding to Akkad. *h-* could be interpreted as an indication of a Common Nostratic origin for the word (with subsequent metathesis in one or another of the language groups, possibly under the influence of tabus), the semantics do not guarantee that the possibility of a borrowing is excluded here (most probably from some intermediate central Eurasian language, though not necessarily Nostratic).

Lexical comparisons are, on the other hand, a rather essential tool for the reconstruction of a proto-language, for they enable us to connect it to a certain period in cultural history and a certain social construct; but, on the other hand, they can be significant only against the backdrop of previously established grammatical correlations which remain the primary argument for the existence of any proto-language, here specifically Nostratic.

In the light of the grammatical arguments cited above (and other such arguments previously employed in studies on Nostratic, cf. Dybo [1978:400-13]) and basing ourselves on strictly linguistic data, we consider Nostratic an historically real proto-language. These data take the lead (both in the case of Nostratic and Indo-European), because a considerable depth of penetration into chronological strata does not, for the moment, permit us to employ extralinguistic data that are essential in, for instance, the corroboration of the existence of relatively late proto-languages such as Common Slavic and Turkic. However, since the set of linguistic arguments that establish the reality of the interpretation of systemic correlations as a proto-language can be applied to the same extent in both the Nostratic and Indo-European proto-languages, as well as in Common Slavic and Turkic, it turns out to be possible—at least in principle—to include all of the above languages in a description of all known languages, that is, to regard them as potential objects for description in a referential, encyclopedic survey ("Languages of the World"), if one is going to attempt to describe as many ancient languages as possible. Of course, in corresponding sections in such a survey, one would have to cite counter-arguments to the postulation of proto-languages. Note, in this regard, Trubetzkoy's (1939) remarks on Common Indo-European. The same thing would also have to be done for questionable modern "languages," such as Macedonian, as well as questionable ancient "languages," such as Pre-Greek Pelasgian.

### The Distinction Between Proto-Languages and Intermediate Stages of Dialectal Evolution: The Problem of Minimizing the Number of Proto-Languages

If the foregoing view is accepted, then one would, in any survey such as "Languages of the World," have to include all the proto-languages whose existence is probable on the basis of internal linguistic arguments, that is, interpretations of systematic correlations *qua* language. As far as chronologically proximal languages are concerned (e.g., Proto-Germanic and Proto-Italic), extralinguistic arguments might also be taken into account. But here a question arises which is essential, both for the theory of comparative historical linguistics and the practical work toward compiling "Languages of the World": how many proto-languages can and/or ought to be posited for the pre-history of every single language in the world? According to Schleicher's *Stammbaumtheorie*, it was assumed that a proto-language (e.g., Indo-European) evolved, via different stages of splits, into individual "sub-proto-languages." Schleicher's theory was supported by many scholars well into the twentieth century, despite persuasive criticism to the contrary on the part of scholars who, armed with data from dialect geography, supported Schmidt's *Wellentheorie*. These data are in line with the conclusions recently achieved by computerized work in historical dialectology, e.g., Wang (1976). Here, it will suffice to cite Sturtevant's Indo-Hittite hypothesis in Indo-European linguistics, an hypothesis supported in several recent publications, cf. Hodge (1974:33), Cowgill (1975), Sturtevant-Hahn (1951). In its modern guise, this theory pre-supposes a break-up of an Indo-European (or "Indo-Hittite" in Sturtevant's terminology) proto-language into two proto-languages: Common Anatolian (Common Hittite-Luwian-Palaic) and the proto-language for all other Indo-European dialects. An exact conceptual parallel is the opposition between Common Semitic and all other languages labeled Hamitic, an opposition that was abandoned long ago in Afro-Asiatic linguistics. The Indo-Hittite hypothesis, based, as it was, on a correct assumption about the relative

antiquity of Hittite vs. other Indo-European languages, runs up against a whole range of counter-arguments which seem irrefutable. First, the verbal system of Hittite, a primary source of support for the Indo-Hittite hypothesis (Cowgill 1975), contained, even in the texts of the Old Kingdom, a number of substantial innovations, innovative even when compared with those in languages, such as Tocharian, which are in many respects even more archaic, albeit documented some three millennia later, cf. Bauer (1976). Second, the use of syntactic elements such as *\*to*, essential both for Sturtevant and some of his more recent followers, (e.g., Hodge 1974), is not specifically Hittite, see Ivanov (1978:192). Third, the preservation of laryngeals in initial position is not only typical for Anatolian, but also Classical Armenian (before Arm. *o* < *\*o* ~ Arm. *a*),<sup>8</sup> as well as (according to Hamp's hypothesis, now supported by additional arguments) Albanian and there presumably also before *\*e* (in particular, the Indo-European argument, which states that, though absent in Anatolian, one can reconstruct *\*he* < Nostr. *\*Hi*, so Illič-Svityč [1971:250]. Therefore, the old form of the Albanian augment in *hε-ngra* 'he ate' can preserve archaic *h* < *\*H*, cf. Petersen (1899:341).

The identification of two intermediate, though not mutually related proto-languages, Common Anatolian and Indo-European (in the classical neogrammarian sense of the term), was precipitated by the desire to preserve the Indo-European proto-language unchanged and was based on a methodologically faulty premise: namely, the realization of a dialect qua dialect by its retentions rather than its innovations.<sup>9</sup> This contradicts both the above mentioned counter-arguments, as well as the fact that Hittite (and other Anatolian languages) belong to a relatively late dialectal group within the Common Anatolian grouping, specifically the Anatolian-Tocharian-Italic group for which the following isoglosses are characteristic: use of *\*k<sup>w</sup>i-* (and not *\*yo-*) as a relative pronoun, presence of mediopassives in *\*-t(o)r/\*-nt(o)r*, all of which are obviously secondary as far as the protolinguistic situation is concerned, cf. Jasanoff (1977). Such dialect groups cannot be viewed as independent proto-languages, but should be delineated in terms of the formal apparatus of linguistic geography, such as

bundles of isoglosses within a continuum of genetically related dialects.<sup>10</sup> The correlation between continuity and discreteness in this field of linguistics permits, in principle, the use of both mathematical terms and devices. This was pointed out as early as 1925 by Hugo Schuchardt, who wrote: "Many years ago I wanted to write an article about 'Linguistics and Mathematics' and publish it in Ostwald's *Annalen der Naturphilosophie*. If I remember correctly, it was about the mathematical representation of genetic relationship (i.e., convergence and divergence) of one dialect with/from other neighboring ones.... It seems to me that, independently of these plans, genetic relations in linguistics can be likened to the mathematical notions of infinity and variability."

As long as a continuum of genetically related dialects is labelled a language, in our case a proto-language, here specifically Proto-Indo-European, and remains unshattered and persists as a continuum, then it is impossible to delineate exact internal boundaries that sufficiently isolate *language* as distinct from *dialect*. This historical linguistic problem is not much different from that of separating a dialect from a language per se in modern dialect geography. For the sake of simplicity, let us cite a very clear example. The prominent linguist and ethnographer, D. K. Zelenin (1927), has distinguished four principal dialect areas in East Slavic: Ukrainian, Belorussian, North Great Russian, and South Great Russian. Today, no one would doubt but that the first two dialects are to be considered separate, discrete languages. This differentiation is rooted in the lengthy linguistic history of the speakers of these languages. Then, too, no one would doubt but that there is an uninterrupted linguistic continuum which begins in North Great Russian and terminates in South Great Russian. This thesis of continuity holds for the *Sprachgefühl* of the speakers of the corresponding dialects, even though there is a certain tendency to create a literature based on one of the dialects, particularly some of the North Great Russian dialects with which some of the most important literary achievements of the last decade were associated. Certainly, V. Belov and N. Rubcov considered themselves authors of dialect literature. They consciously included themselves as authors of Standard Russian, though they retained some traits of their dialects.

The Slavic example was cited because of its unambiguous nature. In this respect, it is different from similar problems that emerge upon an analysis of German and Italian dialects where even more complex problems of differentiation arise in the literary use of dialect. A general principle can be formulated that seems essential both for general linguistic theory and the compilation necessary for "Languages of the World." One has always attempted to minimize the number of "languages" and not to consider as independent languages those dialects that have become severed from a mainstream continuum (note the use of Afrikaans vs. Dutch), or other cases in which specific social and cultural-historical conditions did not conspire to designate a dialect as an independent language. Of course, there are some disputable and borderline cases that emerge when the definition of a dialect as an independent language varies according to concrete historical conditions. Note the differences in the definition of the Latgal situation during the last four decades, e.g., in the early work of Loča, who considered Latgal an independent language. We will not cite further examples from modern dialects and languages, for corresponding borders are here mostly determined by specific linguistic situations. Instead, we are going to discuss the application of the general principle formulated above to the situation of proto-languages.

The principle of minimizing the number of languages generates, as its corollary, the principle of minimizing the number of intermediary proto-languages—beginning with the break-up of a common proto-language, considered as an historical reality rather than as a mere label as in the case of Altaic, through to the emergence of individual, historically attested languages. As a straightforward example, the case of the Common Slavic proto-language may be cited. The emergence, at a certain period in the break-up of Common Slavic, of the Common East Slavic (Old Russian) languages is incontestable. In other words, in the history of any given East Slavic language, three periods can be identified: Common Slavic, Common East Slavic, and historical periods within any given East Slavic language. However, application of this pattern to the South Slavic languages (although the historical and extralinguistic facts of their sufficiently early separation in the

Balkans is beyond any doubt) seems to be quite hypothetical: Old Church Slavic still remains, throughout the early period of its fixation in writing, rather close to Common Slavic. At the same time, it contains features that are specific for this language alone and that distinguish it from other South Slavic dialects. Therefore, terms such as "South Slavic" and "West Slavic," though taxonomically useful, cannot be considered equivalent to certain logically permissible intermediate proto-languages, since there is simply "no time left" for them to have existed subsequent to Common Slavic and prior to the emergence of concrete, historically attested languages (Old Church Slavic, Old Sorbian, Old Czech, Old Polish, Polabian, etc.). Similar reasons compel us to accept the reality of "North Germanic" (Common Scandinavian) as a proto-language, but not the reality of proto-languages that would correspond to the classificatory label "West Germanic."

The above considerations may, it seems, be applied, *mutatis mutandis*, to explain such intermediary dialectal groups as Balto-Slavic, Italo-Celtic, and, possibly, even Indo-Iranian. Although each Indo-European dialect participated in these intermediate dialectal constructs (i.e., for the above three groups respectively: Proto-Baltic and Common Slavic, Proto-Celtic and Latin-Faliscan together with Osco-Umbrian, Mesopotamic-Aryan, Indo Aryan, and Kafiran together with Iranian), it seems exceedingly difficult to prove the existence of actual proto-languages and not just the existence of Indo-European dialect groups characterized and defined by whole bundles of isoglosses which correspond to the classificatory notions implied by the above labels. As for Italo-Celtic, here not only the question of the reality of a corresponding proto-language is unclear, but also the question of the actuality of a Common Italic proto-language (whereas the reality of separate Osco-Umbrian and Latin-Faliscan proto-languages is beyond doubt).

Of particular interest along these lines is the problem of identifying an Indo-Iranian proto-language (that is, Aryan in the strictest sense of the word). The discovery of Mesopotamic Aryan, with traces in Kassitic and Hurrian in the form of loans and proper names attesting to its existence as far back as the beginning of the second millennium B.C., presents a particular difficulty for the establish-

ment of boundaries between the individual Indo-Iranian languages. Probably even at this early date these languages were sufficiently differentiated. Therefore, as a technical term, the compound *aika-wartanna* 'one turn' finds a correspondence in the Iranian languages, cf. the Iranian proper name *\*Varta-aspa* (represented as *Ú-MAR-TA-AS-PA* in late Babylonian cuneiform documents of the 5th century B.C. from Nippur<sup>11</sup>) and Ossetic *æwwærdyn* 'to train a horse', see Bailey (1957). On the other hand, as far as the structure, and not the semantics, are concerned, a compound consisting of a numeral and the verbal stem *\*wert-* finds its parallel not in Iranian, but rather in Skt. *eka-vṛt-* 'only, sole' in the Atharva-Veda (cf. for the Mesopotamian Aryan *tera-wartanna-*, Rg-Vedic *tri-vṛt-* 'thrice turned, rolling on three wheels'). One cannot consider a broader (that is, not specifically equestrian) meaning as completely excluded from Mesopotamian Aryan, especially if the hitherto unexplained Anatolian word for 'new moon' *tanwarta*<sup>12</sup> (for *tan-* cf. Hitt. *dan* 'second'<sup>13</sup>) originates from it. This word is preserved in Old Assyrian documents from the 19th century B.C. as a designation for the ninth month ( $2 \times 4 + 1$ ). On the other hand, the second part of an Indo-Aryan *\*eka-vṛtta* > Assamese *ēta* corresponds semantically as well to the second part of Old Prussian *aina-wārst* 'one time' (with a different suffix in the numeral). This means that this very type of word formation originated within a dialect community which included not only Indo-Iranian, but also (West) Baltic. As for the specifically equestrian significance, it remains typical for Mesopotamian Aryan and Iranian, but not Indic.

A detailed investigation of the Kafiran lexicon has shown, in principle at least, considerable proximity between these languages and Indo-Aryan. However, the assumption of a natural proto-language for both Indo-Aryan and Kafiran is hampered both by the known differences in phonological evolution (where Indo-Aryan is, in some respects—particularly so far as reflexes of *s* in its phonological relation to a palatalized series are concerned—closer to Iranian than to Kafiran) and by lexical divergences. Such Kafiran lexemes as Proto-Su. . . *čūū* 'curdled milk', Kati *trū*, *trū* find exact correspondences in Iranian (cf. Av. *tūiri* id, see Fussman (1972:59). Especially telling are forms such as Kati *dić* < *\*diju* 'tongue' vs. Av. *hizū*,

Vedic *juhū*, see Fussman (1972:441). It is evident that we have to identify Kafiran as a separate protolinguistic entity alongside Mesopotamic Aryan, Common Iranian, and Indo-Iranian. It is not altogether impossible that dialects of Indo-Iranian, for which traces can be found in several North Caucasian loans, were close to Kafiran (particularly if one considers reflexes of palatalized stops of the type glottalized c). Cf. forms of the type Common Daghestanian *\*mYǝVr*, *\*mVč'Vr* 'beard' (see Gigineisili 1977:135) vs. Skt. *śmaśru* and in non-palatalized forms like Alb. *mjekrë*, Lith. *smakra* in more westerly *satem*-dialects contiguous with varieties of Hittite with *zamankur*, whereas Hittite proper normally shifts *\*-ku-* > *-šu-*. If one wants to presume the existence of a Common Aryan proto-language that includes all of these dialects, then one must push the timeframe for the existence of an Armenian-Greek-Indo-Iranian community (which gave rise to all of these dialects) even further back in time. However, if one takes the results of recent research into account, then this same dialect community is to be regarded, not as a proto-language, but rather as a group of dialects of the Indo-European proto-language that are sufficiently close to other Indo-European dialects; features typical for this community, particularly the verbal augment, can be revealed in such flanking dialects as Albanian in the West and Tocharian in the East.

In a survey such as "Languages of the World," it seems reasonable to include only articles of the following type: "Indo-European Proto-Language," where data on early dialect divergences of this "language" would have to be detailed, including the groups of dialects listed above; "The Baltic Proto-Language," where its close affinity with Common Slavic has to be indicated; "The Common Slavic Proto-Language," where its affinity to common Baltic has to be indicated, etc. In other words, intermediate groups would have to be reflected in the survey, but not in separate entries devoted to them.

This procedure would also be valid for those proto-languages that are chronologically most remote, though there would be some basis for including the West Nostratic proto-languages (Indo-European, Kartvelian, Semito-Hamitic) as one dialectal group which differs from the East Nostratic proto-languages. Nevertheless,

we are here dealing with dialect divergences within the Nostratic proto-language itself, not with any actual intermediate proto-languages. One should insist on such a procedure, for sometimes scholars attempt to restructure the Nostratic hypothesis along the lines of a *Stammbaum*-pattern, a pattern which presupposes sequential splits of a Nostratic proto-language into intermediate proto-languages; such a configuration is sometimes cited as an argument against Nostratic theory itself. It is clear that for such types of splits an exceedingly long period of time would be required; and this would come as a crass contradiction of the real chronology of the history of mankind. Now, the problem of Altaic is recast in a new light. Languages which have been united under the convenient classificatory label "Altaic" (Turkic, Mongolian, Tungus-Manchu, Proto-Korean, and, probably, Proto-Japanese) would emerge from a certain portion of the (East) Nostratic dialectal continuum without an intermediate Altaic proto-language. As for Uralic (Proto-Finno-Ugro-Samoyedic), its linguistic reality is much more definite than that of a Common Finno-Ugric proto-language, which is assuredly indicated by the facts mentioned above on Uralic verb morphology. It would seem useful to reflect these alternative viewpoints in sections with titles such as "The Nostratic Proto-Language," "The Uralic Proto-Language," as well as in more specific articles devoted to the proto-languages of separate groups of Finno-Ugric languages such as "The Ugric Proto-Language," "The Permian Proto-Language," etc. One has to attempt to select the minimal number of proto-languages required to describe all the chronologically intermediate groups of dialects.

It seems reasonable to employ a similar procedure when describing the languages of Africa. The principal proto-languages of microfamilies (according to Greenberg's amended classification) would have to be given, but not the intermediate proto-languages; the same would hold for the Amerindian languages whose principal proto-languages of their macrofamilies (as already noted by Sapir and others) would have to be given. Though numerous doubts have been voiced about E. Matison's Amerindian hypothesis,<sup>14</sup> as well as Greenberg's Indo-Pacific hypothesis, the existence of correspondingly archaic proto-languages

seems sufficiently plausible as to be included in the list of proto-languages to be described in the referential, encyclopedic survey (with, of course, the same historical-bibliographical reservations as are to be required in the article, "The Nostratic Proto-Language"). On the same grounds, an article such as "The Common Australic Proto-Language," in accordance with the recent discoveries of Capell, will be needed, though further dialectal groupings within this proto-language might not correspond to any "actual" proto-languages. A similar approach would presuppose a sufficiently economical description of both Austronesian languages (where the existence of an Austronesian proto-language is beyond doubt, though there are probably no such intermediary proto-languages as North Austronesian) and the Austro-Asiatic languages.

### **The Descriptive Strength of a Proto-Language**

The inclusion of principal proto-languages in "The Languages of the World" (as well as the consistent elimination of intermediate ones according to the principle of minimization), is not just stipulated by requirements for an exhaustive corpus of old languages in a survey of this kind. The following arguments seem even more compelling: first, in many cases the proto-languages are better known than some individual ancient languages preserved in fragmentary inscriptions that are often badly and arbitrarily interpreted, in proper names, in glosses, etc. To make matters simpler, let us take a very clear-cut example. It would seem obligatory to include the language of the early futhark runic inscriptions in the roster of the older Germanic languages. Equally obligatory, under the entry "The Common Germanic Proto-Language," would be an exhaustive characterization of the phonological, grammatical, and lexical structures of this "language"; whereas under the rubric of "The Language of the Later Futhark Runic Inscriptions," it would be sufficient to note that this "language" preserved a number of phonological and grammatical features of Common Germanic (followed by a list of principal differences). In general, the description of the language of the older futhark inscriptions will be less rich than that of Common Germanic. The same

would hold, for example, for Ibero-Celtic and the Common Celtic proto-languages.

Among the older Indo-European languages there are many that are known only by very fragmentary, mostly toponymic, exceedingly terse and formulaic inscriptions. Research on the scanty remnants of Thracian, Illyrian, and other feebly attested Indo-European languages has been quite intensive, so much so that the number of studies treating these *rarissime* has by now far surpassed the very number of fragments. In general, this situation has not made the structure of these languages any clearer, so that the corresponding entries for "The Languages of the World" will not give a complete description of their linguistic structure. The reader will obtain a lot more information from a sentence or two in each of these entries in which the origin of the language in question from within a certain dialect of the Indo-European proto-language is described. No doubt the Common Indo-European Proto-Language has been studied in much greater detail; and its description for the later, immediately pre-dialectal period will be much fuller (provided unnecessary hypercriticism is avoided) than the entries dealing with archaic languages known only from very scanty data. More than that: these very entries will be based, to a large extent, on conclusions about the structure of the Indo-European proto-language.

What is said here also applies to the extinct Yeniseyan languages known only from the sketchy notes of travelers in the 18th century, as opposed to a rather complete characterization of the Yeniseyan proto-language which can be given on the basis of a comparison of Ket, Yughan, and the extinct Kottic language. (As for Yughan, the above thesis concerning the difficulties inherent in making a distinction between *language* vs. *dialect* can be applied. Until recently this almost totally extinct language was considered a dialect of Ket, namely, Symic Ket, despite the fact that there is no mutual intelligibility between these languages without the aid of translation; and their lexical and phonological structures differ considerably.) By way of example, one can cite the possibility of reconstructing original strings of affixes, later remodelled in each of the historically attested languages (cf. OS order in Celtic vs. SO order with both transitive verbs and verbs

of the type Ketic *d-a-r'* -*æ<sub>n</sub>-dij* 'we will arrive', lit. 'we-for-ourselves-will-arrive', and so on).

Another lucid example of the explanatory power obtained by the inclusion of proto-languages, as opposed to the inclusion of language families represented by older languages, is the inclusion of Hurrian-Uralic as a general notion by means of which both individual Hurrian dialects (spoken during virtually an entire millennium in different areas of the Near East) and Uralic (sometimes referred to as a Hurrian dialect, though by the term "Hurrian" the Hurrito-Urartian proto-language is usually intended) can be explained. Concerning the rather substantial degree of Hurrito-Urartian proto-language divergence in some vocabulary that, in the final analysis, is of common origin, one can make inferences on the basis of items such as Hurr. *šaphali* vs. Urart. *salmathi* 'left, left-handed' (in the latter there might be contamination with Semitic forms akin to Akkad. *šumēlu*). The presence of explicit functional differences between forms originating from some one proto-language makes references to it rather useful when explaining each proto-language separately (as well as when exploring the possibilities of relationships with other languages). In line with this, the Urartia indefinite pronoun *-ija*, *-je* (see Diakonoff 1971, Hass-Wilhelm 1974:92) can ultimately be explained as a protolinguistic element which corresponds to Nostratic *\*ja*, supposedly with the original value of an interrogative pronoun which later developed into a relative pronoun in Indo-European, Semito-Hamitic, and Finno-Ugric, where indefinite significance is also presumed. Similarly, in discovering the history of the Hurrian-Urartian ergative marker (Hurr. *š*, Urart. *še*) a description of the protoform is useful, a form that can be compared with a Nostratic pronominal and by means of which the primordial Indo-European active case in *\*s* can be explained.

Next, the inclusion of entries covering all principal proto-languages in "Languages of the World" will permit one to substantially clarify the description of individual languages. Since such "languages" as Turkic, Bantu, or Slavic themselves incorporate rather similar languages, it is not necessary, in every entry depicting individual languages, to reiterate those characteristics which were inherited from the proto-language (i.e., Proto-Turkic,

Proto-Bantu, Proto-Slavic), for they can be given in an entry describing the individual proto-languages. On the other hand, the description of a language which has become significantly different in grammatical structure from its corresponding proto-language (such as Modern Bulgarian vs. Proto-Slavic) will become more distinct precisely because of a contrastive comparison with its protolinguistic structure.

Next, when reconstructing protolinguistic structures, one succeeds in revealing such characteristic features of a language type that, in historically attested languages, appear in less distinct form. As the evolution of a language in time can be described from the perspective of information theory as the transmission of a message through a communication channel with noise, so too the reconstruction of a proto-language can be considered as the restoration of an original message. It is clear that, when compared with individual historically attested languages, certain fundamental structural features appear in a more evident shape in the original message. As an example in support of this contention, one can cite the discovery of traces of languages with "active" constructions (when distinguishing between "volitional" and "non-volitional" action) in Indo-European. This was possible only when the protolinguistic situation had been reconstructed.<sup>15</sup>

The above example of normal declension in Tocharian convincingly shows that the revelation of Tocharian's connection with the archaic heteroclitic type in Indo-European becomes possible only by reconstruction of the common Tocharian proto-language. In both of the Tocharian dialects attested in written documents, this archaic system already presents itself in a form complicated by more recently evolved gender classification superimposed on the original Indo-European opposition between animate and inanimate noun classes. As far as the chronologically less remote languages are concerned, one may observe that the discernment of the typical features of a language with open syllables of the type represented by Modern Japanese<sup>16</sup> appears exactly possible for Proto-Slavic. At the same time, this comparison appears very useful for understanding concrete phenomena in such individualized languages of Slavic as Old Church Slavic, see Avram (1978:24-5).

In accordance with our comprehension of linguistics as a study of relations among language systems, one can take as a given that a successful description depends on the correct choice of a system to serve as the linguistic *étalon* of this description. The greater the explanatory power of the *étalon*, the more effective the description. As far as the aims of an encyclopedic survey such as "The Languages of the World" are concerned, the most promising *étalon* for descriptive purposes is the proto-language of any given family.

### NOTES

#### Ivanov (1980): Proto-Languages as Objects

1(E). The reference here is to Illič-Svityč (1963). In 1983 a revised list of reconstructed Indo-European lexemes borrowed from Proto-Semitic was compiled by A. Dolgopolsky and is available from that scholar.

2(E). Cf., most recently, Bomhard (1984). Unfortunately, however, as pointed at at some length in the Foreword to this collection, even this most recent effort is flawed by too many implausible or simply incorrect comparisons.

3. On the comparison of Kartvelian and IE *\*-m-*, see Illič-Svityč (1976:64).

4. See Klimov (1977), McLendon (1978:4-5), and, for comparisons with Indo-European, Lyons (1968:293).

5. For details on these semantic relations, see Otten (1973:36).

6. See the review of this material by Ivanov (1974:192-5).

7. See *Tungusian* (1977:92). Here (p. 477), there is an error in the description of the semantics of the Manchu word. It is Manchu that compels us to assume that the significance 'rat' (even if it belongs with 'mouse') was present in Common Tungusian.

8(E). Absence of Arm. *h* (< IE [x]) before *e* is probably due to the fact that Nostratic and Indo-European "strong" laryngeals rarely appeared before *\*e*. In Armenian, the voiceless "strong" laryngeal of the proto-language remained as *h-* in word-initial position, whereas voiced "strong" laryngeal (*\*γ-*) disappeared, together with the "weak"

laryngeal(s). In Hittite, both  $*x-$  and  $*\gamma-$  became  $*h-$  (both before  $e < *e$  and  $*a < *a, *o$ ). In itself, preservation of Indo-European laryngeals, not only in Hittite but also in Armenian (and, possibly, Albanian), cannot be regarded as an argument against the Indo-Hittite hypothesis.

9(E). One of the pioneers of Nostratic studies, A. Dolgopolsky, also supports the Indo-Hittite hypothesis. His main argument is the absence of the feminine in Hittite-Luwian. This absence has to be regarded as a Hittite-Luwian archaism inherited from Nostratic, whereas the presence of the feminine in the other "branch" of Indo-European (that is, in all other groups except Hittite-Luwian) is an innovation.

10(E). Bundles of isoglosses of this kind do not necessarily speak against the Indo-Hittite hypothesis: mutual borrowings between or among dialects often come easily, and hence there is the possibility that, as an Indo-European dialect, Hittite-Luwian acquired several innovations from other Indo-European dialects at some very early stage.

11. See Zadok (1975:247), though the comparison with Mesopotamian Aryan is missing here.

12. This reading is justifiable as per Matouš (1978:No. 3.5.218, fn. 6).

13(E). Cf. Lyc. *dda-qasa* 'second son', *dde-newele* 'second offspring', etc. (*dd-* < IE  $*d-$ ).

14(E). A recent overview of one interpretation of the genetic affiliations of (some) Amerindian languages is summarized in the Foreword to this collection.

15. See Klimov (1977) and the research history and literature presented in this volume.

16. See Polivanov (1928) and Martinet (1978) on this typological similarity.

A PROBABILISTIC HYPOTHESIS CONCERNING THE  
OLDEST RELATIONSHIPS AMONG THE LANGUAGE  
FAMILIES IN NORTHERN EURASIA

*Aaron B. Dolgopolsky*

**Introductory Note**

Inquiry into distant linguistic relationships usually consists of three investigative stages:

1. A preliminary assessment of the advisability of making the comparison. Quite simply, one has to question what languages, or language families, or phyla are worth comparing. Toward this end, one examines the basic, core lexemes and grammatical morphemes, but initially only at a superficial, non-etymological level, for etymology is based on diachronic phonological rules, and these remain to be uncovered.

2. An identification of phonological correspondence and of the rules of diachronic phonology operative in the languages being compared as they apply to essential roots and affixes.

3. An etymological analysis of certain lexemes, roots, and grammatical markers.

The following article was written over two decades ago and addresses itself only to the first of the above three stages; it is concerned with a preliminary evaluation of the advisability of comparing certain languages. Soon after it was written, however, work was begun by Illič-Svityč and myself that was directed toward identifying phonetic/phonological correspondences and, thereafter, etymological comparisons based on these phonetic/phonological correspondences, cf., for example, Illič-Svityč (1971) as a sample of these efforts. Work along these lines is still in progress.

It is now obvious that the diachronic phonology achieved from these efforts invalidates several of the correspondences that were proposed over two decades ago. For instance, we now know that an IE *\*duō* 'two' is to be compared with a Semitic *\*tʾm* 'twins', not with Semitic *\*θnǝj* 'two' (Egyptian *śn*, Berber *śn*). On the other hand, a comparison of IE *\*duō*, the Tungusian prototype for 'two', and Korean *tu-* proved correct. Then, too, we must now

abandon comparison of IE *\*okw-* 'eye': Sem. *\*sayn-* and Cush. *\*sil-* 'eye'. Subsequent etymological investigation showed that the Hamito-Semitic congener of an IE *\*okw-* is different, namely, *\*ʕVkʷ-* 'see' → 'know', cf. Illič-Svityč (1971:Entry #118), Dolgopolsky (1973). Today, we know that IE *\*kerd-* 'heart' corresponds to Kartv. *\*m-k'erd-* 'chest', Hamito-Semitic : Chadic *\*k'Vrd* id. and not to Sem. *\*k'VlVb-*, Kartv. *\*gul-*, Mong. *žürüken*, etc., cf. Illič-Svityč (1971:324-5, Entry #200).

Nevertheless, most of the other "early" correspondences proposed here, including those established for the pronouns, have later proven to be etymologically correct.

In light of these remarks, I ask readers not to regard the following article as an etymological one, but rather as a methodological one, a preliminary attempt to establish guidelines by which one may evaluate the advisability and possibility, if not the efficacy, of comparisons and remote genetic affiliations. As for etymological studies, accomplished with increasing sophistication and insight, the interested reader can now turn to more recent studies by younger scholars who have followed in the footsteps of Illič-Svityč as a dramatic pioneer.

A. Dolgopolsky, August 25th, 1983

### A Probabilistic Hypothesis Concerning the Oldest Relationships Among the Language Families of Northern Eurasia

Approximately sixty years ago, the eminent Danish linguist, Holger Pedersen, proposed a hypothesis concerning a possible genetic relationship between Indo-European, Semito-Hamitic, Uralic, Altaic, and some other languages.<sup>1</sup> Until now, this hypothesis, which I have provisionally labelled the "Sibero-European Hypothesis," has been neither proven nor disproven.<sup>2</sup>

A primary difficulty in resolving this issue resides in the lack of a mathematically rigorous procedure to substantiate linguistic relationships. When dealing with closely related languages whose genetic affinity is rather obvious, as for example, in the case of Turkic vs. Indo-European, we may dispense with such rigorous proofs. It is quite another matter, however, with distance linguistic

relationships. Here, we require mathematically rigorous methodologies which will permit us to distinguish between languages that are actually genetically related and those that are merely fortuitously similar.

A phonetic similarity between semantically like morphemes in different languages can have three conceivable explanations:

- 1) pure coincidence
- 2) borrowing
- 3) genetic affinity.<sup>3</sup>

If there are morphemic coincidences between specific languages and one can prove that fortuitous coincidence is highly likely (and simultaneously demonstrate that borrowing is highly unlikely), then the only plausible explanation for such coincidences is genetic relationship.

In this paper I will propose a procedure whereby one can prove a putative genetic relationship between languages. This procedure is based in establishing a probability calculus for the fortuitous phonetic coincidence of morphemes where the degree of such fortuitous coincidence is not less than that displayed by (the overall relationship between) the languages in question.

To make language comparison more testable and probative, two approaches are employed:

1) Comparison of several languages. Not only two languages or two language families, but more than two languages or language families are to be compared. When only two languages are being compared, then the probability of purely fortuitous phonetic coincidence in semantically identical morphemes is very high. However, when such morphemes are compared in four or five languages, coincidental probability is quite low. The probabilistic chance of pure coincidence of one morpheme compared in five languages is less than when dozens or even hundreds of morphemes are compared in but two languages, see Greenberg (1953:271-2).

The majority of linguists who have investigated the possibility of remote relationships among Sibero-European languages restricted themselves to the comparison of but two language families; this is the basis of the Indo-European - Semitic hypothesis of Hermann Møller and

Antoine Cuny, of the Indo-European - Uralic hypothesis of Björn Collinder, the Uralo-Altaic hypothesis of Matti Räsänen and others, the Kartvelo-Sumerian hypothesis of M. Tsereteli, the Indo-European - Sumerian hypothesis of Sch. Otran, the Uralo-Chukchee hypothesis of J. Angere and K. Bouda, and so on. It seems that a simultaneous comparison of all the Siberian languages would be more effective.

In this paper, seven language families will be compared: Indo-European, Hamito-Semitic, Uralic, Altaic, Chukchee-Kamchatkan, and the Sumerian languages.<sup>4</sup>

2) Statistical selection of semantic values represented by morphemes which are relatively impervious to change. Another approach which makes language comparison more effective consists of a carefully restricted data base of what is to be compared; only those morphemes are considered that are characterized by a high degree of stability in any or all of the languages compared.

With the use of statistical procedures, one can make an inventory of those semantic values least impervious to change; and one can then compare morphemes with these semantic values in different languages. As soon as a phonetic similarity is found, then the probability of purely fortuitous coincidence is calculated, a probability which should not, however, be less than the statistical mean. If this probability is very low and if the phonetically similar morphemes have a semantic value that characterizes grammatical markers and sectors of the lexicon that are not normally subject to borrowing (where these are not members of the following categories: onomatopoeia, interjections, and *kindersprache*), then a genetic relationship can be considered as proven.

The procedure for validating linguistic relationships consists of three steps:

STEP ONE. Uncover the extent of the "historical" stability of morphemes which have different semantic values; in other words, discover what semantic values of morphemes remain unchanged for long periods of time and in what semantic guises they are rapidly replaced by other morphemes. Note, for example, that roots with the value "star" are more stable than those with the value "lightning." Throughout the history of Romance, "star" has always been conveyed by the same radical (cf. Lat.

*stella*, Fr. *étoile*). As for "lightning" in the same group of languages, there were no less than three replacements. As a result of these replacements, four roots currently represent "lightning" in Romance: 1) Rum. *fulger* < Lat. *fulgur*; 2) Sp. *relampago*, Port. *relampago*, Cat. *llampec*, It. *lampo*; 3) Fr. *éclair*, 4) Rhaeto-Romance (Romansch) *sajetta*. Note, further, "star" vs. "lightning" in other language families and phyla. In the Germanic languages (twelve were investigated), there is no replacement for "star," but two for "lightning." In Slavic (ten languages were investigated), there is one replacement for "star" and three for "lightning." In Indo-Iranian (twenty-one languages were investigated), there is no replacement at all for "star" vs. five replacements for "lightning." The older stages of Indo-European (from Proto-Indo-European to the languages of the individual groups, eleven languages in all), there is one replacement for "star" and ten for "lightning." In Semitic (four languages investigated), there is one replacement for "star" and no replacement for "lightning." In Dravidian (with four languages investigated), there were no replacements at all. In Finno-Ugric (four languages investigated), there were two replacements for "star" and four for "lightning." In Daghestanian (seven languages investigated), there was one replacement for "star" and two for "lightning." In Turkic (fifteen languages investigated), there was no replacement for "star" and six for "lightning." For the total of the above languages, there were six replacements for "star" vs. thirty-six replacements for "lightning."

In establishing degrees of morphemic stability, the following procedure is considered optimal. The entire history of each language is to be investigated, and for each the number of times its morphemes with a given meaning have been replaced is to be determined. The replacement figures for a given meaning in all languages are to be totalled, and this yields the general quantity of morphemic replacement for a given meaning in all the languages considered. This figure could be considered an index of the degree of morphemic replacement: the higher the index, the greater the propensity for replacement of a given semantic value (e.g., "star" vs. "lightning").

However, such an "optimal" procedure is practically impossible because the history of various languages is

not always retrievable (from written records, etc.). Therefore, we must substitute this procedure with another, a procedure that is feasible. We consider groups of related languages and then discover how many different roots are represented in a given group with a given meaning. In each language, only the principal synonym is considered. If we find the same morpheme in all the languages of a given group, then this would imply that there were probably no replacements of this morpheme in the history of this language group. If we find  $n$ -roots, then this means that there were no less than  $n-1$  replacements. The figures that indicate one given meaning in all the languages are then totalled. A sum obtained by this means is the index of morphemic replacement for the given semantic value (e.g., "star"). In terms of this procedure, some replacements could elude consideration if they left no traces in both contemporary and historical (written) sources. Therefore, these figures are not to be viewed as an absolute indication of the exact quantity of replacements in a linguistic history, but rather as the indices of the degree of propensity of a specific semantic value (e.g., "star" vs. "lightning") for morphemic replacement.

Despite these and other shortcomings, the procedure proposed here corresponds better to our avowed goals than a calculation of replacements on the basis of written sources in those few languages in which these are available. This latter procedure has, of course, been employed by Robert B. Lees.<sup>5</sup> Understandably, Lees could only calculate correlative data for thirteen languages for which the entire history is something less than 15,000 years. It is impossible to draw statistically valid conclusions from such restricted data. The methodology proposed here not only enables one to take the rather short history of written languages into consideration, but also multi-millennial historical periods from a proto-language (of any group) forward to the status of the contemporary language. I have investigated 140 languages (64 Indo-European, 4 Dravidian, 12 Daghestanian, 4 Abkhazo-Adyghan, 4 Kartvelian, 7 Semitic, 17 Turkic, 6 Mongolian, 6 Tungusian, 10 Finno-Ugric, and 6 Samoyed). When considered in their entirety, their histories would cover some 200-300 thousand years. Data from the aboriginal languages of Australia (159 languages according to

W. V. Schmidt), Africa, and America have also been looked at, but were not included in the general statistics, as semantic data significant for our purposes here are lacking for these languages (e.g., terms for 'heart', 'louse', 'horn', and so on). It is important to note, however, that even in these languages more or less the same semantic values were characterized by the same high degrees of morphemic stability as was the case for the languages of Europe and Asia. Some categorical differences of this kind were taken into account here when the following list of semantic values for determining genetic relationships was compiled (see below).

Clearly, we are interested in listing the semantic values of only the stablest morphemes. For this purpose, all those semantic values in Carl Darling Buck's (1949) dictionary of synonyms were considered—specifically, values that have no more than ten replacements, (that is, ten divergent forms out of the total inventory of the thirty-one Indo-European languages examined by Buck). Moreover, all semantic values with no less than a 60 percent stability factor were excerpted from the list by Lees, see Swadesh (1955:Table 2). Finally, some semantic values that appeared stable and therefore worth checking out were added to those in Buck and Lees. Altogether, 250 semantic values were examined.

As a result, a list of semantic values was compiled in which all values were ranked according to their degrees of morphemic stability. The list is headed by the following semantic values:

- 1) no replacements: the numerals '5', '3', '4', '6'; the first person pronoun.
- 2)  $1-1\frac{1}{2}$  replacements:<sup>6</sup> the numerals '2', '7', '8'.
- 3)  $2-2\frac{1}{2}$  replacements: second person singular pronoun.
- 4)  $3-3\frac{1}{2}$  replacements: 'who.'
- 5)  $4-4\frac{1}{2}$  replacements: the numeral '10', first person plural pronoun, the numeral '1', 'tongue', second person plural pronoun, the numeral '9'.
- 6)  $5-5\frac{1}{2}$  replacements: 'name.'
- 7)  $6-6\frac{1}{2}$  replacements: '100', 'what.'
- 8)  $7-7\frac{1}{2}$  replacements: 'eye', '20', heart.'
- 9)  $8-8\frac{1}{2}$  replacements: 'tooth', prohibitive NEG ('no, non-, not'); verbal NEG, that is, 'don't sleep!' vs. 'I don't sleep; respectively, 'nit' ('louse', larva or

egg).<sup>7</sup>

- 10) 9-9½ replacements: 'finger/toe nail', 'louse',  
'new moon, crescent of the moon',  
'tear.'
- 11) 10-10½ replacements: 'water', 'dead', 'hand.'
- 12) 11-11½ replacements: 'night', 'blood.'
- 13) 12-12½ replacements: 'horn', 'full', 'sun', 'ear', 'salt'.

To avoid double counting of the same root, (that is, to observe constraints on the mutual inter-dependency of linguistic events [necessary for the proper calculation of the probability of fortuitous coincidences]), those semantic values that are often represented by derivatives from roots which express or represent other semantic values also present in this list must be eliminated from the above inventory. We therefore eliminate the semantic value '20' since it is often a derivative of '2'; also 'we' since it is often derived from the root present in 'I'; and 'you' plural since it is often derived from the root for 'thou.' Note that one and the same morpheme is often both a negator of indicatives and a prohibitive particle (cf. Russ. *ne*). The word for 'nit' often contains the root(s) present in the word for 'louse' and is therefore eliminated. Therefore, '20', 'nit', 1st pers. pl., 2nd pers. pl. have been eliminated from the list. 'Who' and 'what' are considered as one. The distinction between prohibitive NEG and verbal NEG has also been eliminated for the same reasons.

Having taken the aboriginal languages of Australia into consideration, the following adjustments have been made: 1) the semantic values for 'winter' and 'new moon' have been eliminated, as they are lacking in many languages of the world; 2) the numerals '3' through '10', as well as '100', have been removed from the list, since they do not have specific roots in many Australian languages, the languages of the Americas, and so on; 3) '1' has also been removed since its stability is rather weak in many Australian languages.

Now, extrapolating from this "adjusted" list, the first fifteen values for the final inventory for the study of the relationship between languages and language families appear as follows:

- 1) first person marker
- 2) '2'
- 3) second person marker

- 4) 'who', 'what'
- 5) 'tongue'
- 6) 'name'
- 7) 'eye'
- 8) 'heart'
- 9) 'tooth'
- 10) verbal NEG (both negative proper and prohibitive)
- 11) 'finger/toe nail'
- 12) 'louse'
- 13) 'tear' (noun)
- 14) 'water'
- 15) 'dead'

STEP TWO. On the basis of phonotactic statistics for changes in the world's languages, we can list the most usual (and probable) phonetic correspondences in diachronic change. For example, initial *m*- remains unchanged 90 percent of the time. The change *b* > *m* or *w* > *m* is highly unlikely, and other possible sources of *m* are even less likely. A refined statistical analysis of these sorts of properties has not yet been accomplished. Here, speech sounds have been divided into several types and thereby distinguished in such a way that phonetic correspondences inside a "type" are more regular than those between different "types." Therefore, if we say that *t*, *d*, *dh*, *t*, *θ* belong to one type, this implies that the shifts *t* < *d*, *d* < *t*, *θ* < *t*, and the like occur more often than shifts such as *t* < *b*, *d* < *k*, and so forth. The following preliminary grouping of consonants into type categories has been proposed: 1) Type *P*: labial obstruents (*p*, *b*, *f*); 2) Type *T*: dental obstruents (except *s*, *š*, *z*, *ž*); 3) Type *S*: *s*, *z*, *š*, *ž*; 4) Type *K*: velar and post-velar obstruents (*k*, *g*, *χ*) and affricates such as *c*, *3*, *c*, *3*; 5) Type *M*: *m*; 6) Type *N*: *n*, *ñ*, and non-initial *ŋ*; 7) Type *R*: *r*, *l*; 8) Type *W*: *w* and initial *u*-; 9) Type *J*: *j*; 10) Type *Ø*: laryngeals, *Ø*-consonants, and initial *ŋ*; Vowels are not considered as they are regarded as too unstable; they display widely divergent change patterns. Such a division into types based on intuitions about probabilities of phonetic change (well known to all comparatists) can, of course, be imprecise; and this problem will be taken up after a statistical investigation.

STEP THREE. We now proceed to investigate languages whose relationship we intend to establish. For this purpose,

we will utilize all fifteen semantic values from our final list. From among the morphemes expressing a given semantic value, only the earliest attested are considered in every language family (if it is impossible to determine which is *the* oldest, then several of what are presumed to be the oldest morphemes are taken). For determination of the oldest morphemes, standardized procedures are used. The degree of "morphemic spread" (i.e., extension of grammatical categories: N, N and V, V and Adj) is considered, internal reconstruction applied, and so forth—all, in fact, of the normal comparative procedures. In examining data in this fashion, we also stipulate phonological similarities and calculate the probability of fortuitous coincidence, which is in no case less than the statistical mean. We consider those sounds to be similar that belong to any one of the above types (Types 1-10), see STEP TWO.

Now, let us compare Indo-European, Semito-Hamitic, Uralic, Altaic, Chukchee-Kamchatkan, Kartvelian, and Semitic in light of the fifteen categories presented in the above list.

The result of this comparison is schematicized in Table A in which morphemes are transcribed by means of the phonetic types (1-10). To make this comparison clearer, sounds which coincide in different languages are transcribed by capital letters. Failure to enter a morpheme in a given box signifies that morphemes with the given meaning differ in different languages (in other words, there is nothing to compare). Morphemes in parentheses are not calculated. The final right-hand column contains the probability  $P^i$  of a fortuitous coincidence that is not less than the coincidence observed for the given semantic value  $i$ . The probability coefficient  $P^i$  is rounded off to the next highest standard deviation.

Table A gives the phonetic similarities (as outlined by phonetic type) for the following morphemes:

1) FIRST PERSON: verbal desinences of the first person singular *\*-mi/-m*, first person singular pronoun (in all oblique cases) *\*(e)m*, etc. The nominative *\*eq(h)ōm* has, it appears, resulted from the addition of deictic elements to the radical *\*m*, which happened to be placed at the end of a highly frequent word; this root has been phonetically eliminated in several languages. Possibly

an older subject pronoun type *\*mi* preceded *\*e(q)hōm*, a type *\*mi* preserved in the verbal endings *\*-mi/-m*. // Uralic: *\*mynd* ~ *\*mind* 'I'.<sup>9</sup>; Finn. *mind*, Mord. *mon*, etc. / Samoyed 'I': Nenets *man*, Nganasan *mannang*, Selkup *man*, *mat*, Kamasin *man*, etc. / Yukaghir *met* 'I', Chuvantzy *mota* 'I' // Altaic *\*min* 'I' (so Ramstedt 1957:79) is preserved in Turkic *mñn-bñn*, Literary Mongolian *bi* (gen. of *min*), Gold *mi* (*bi*), obl. case *min-*, Evenki, Lamut, Ude, Manchu *bi* (obl. case *min-*) // Chukchee-Kamchatkan: Itelmen *ki-ma*, *ki-m*, *k(i)me*, 'I', *ma* 'me' (dat.), *muza* 'we', Chukchee *gə-m*, 'I', *muri* 'we', Koryak *g-mmo* 'I' *muju* 'we' // Kartvelian: Georgian *me*, Chan *ma(n)*, Svan *mi* 'I'.<sup>10</sup> // Sumerian *ma-e*, *ma* 'I'.<sup>11</sup>

We are unable to reconstruct the proto-morpheme for the Semito-Hamitic languages.

2) "TWO"

T: Indo-European Gk. *duō*, Lat. *duo*, Skt. *dva-*, etc. // Sem. -Ham.: Sem.: Ar. *ṯn*, Ara. *tr*, Hebr. *šn*, etc. / Berb. *sin* / Egyp. *sn*. The proto-radical could be TN (*ṯn*, *tn*) or SN // Altaic: Gold *dūer*, Korean *tuur*, *tuir*, etc.<sup>12</sup>

KT: Sem-Ham.: Cushitic: Kaffa and Mocho *gutta* / Chadic: Kuseri and Muzgum *hudyu* // Uralic: Finno-Ugric: Fin. *kaksi* < *kakt* + *i*, Erzja *kavto*, Ostyak *kāt*, Hung. *két*, etc. / Samoyedic: Motor *kydy*, Taigi *kidde*, Nenets *sidā*, Nganasan *siti*, Selkup, *sede*, *šite*, etc. Enets *sire*, *side*, Kamasin *šidə* (ST < KT) / probably also North Yukaghir *ki(d)*, Omok *kit*, Chuvantzy *kuen*.

It is not altogether impossible that Sem-Ham. radicals of the type ST (Cushitic: Arbore *sada*; Chadic: Bharein *siidi*, Margi *sidai*, Molgoy *šida*) also originated from Sem.-Ham. KT as a result of palatalization of the initial consonant. Since the Sem.-Ham. radical ST is highly restricted in its distribution across languages, it has not been calculated.

K. . . : Altaic: Mong. *qoyar* '2', *qorin* '20' // Chukchee-Kamchatkan: South Kamchadal *qaz*, West Kamchadal *qasx* '2' < *\*qad-* or *\*qaz-*, Chukchee *kuvrdt* 'pair' is of the same origin. Chukchee *nirāk* 'two' possibly changed its initial consonant as a result of adding an initial prefix on the analogy of Chukchee *nərok* '3' and *nərak* '4'. Koryak *nəccāx* 'two' is of the same origin (Chukchee *-p-* and Koryak *-č-* < *\*d* or *\*z*).<sup>13</sup>

Kartvelian and Sumerian numerals, as well as the

Kushitic numerals of the type RM (Galla *lâma*, Afar *lammei*, etc.) have no parallels in other Sibero-European languages.

### 3) SECOND PERSON

T: Indo-European *\*tū*, *\*t(e)w-*, *\*t-* 'thou.' // Sem.-Ham.: Sem. verbal prefix of the 2nd pers. *t-*, Arabic *'an-ta* 'thou', Hebrew *'attā* 'thou', etc. / Cushitic: Sidamo *ate*, *ati* 'thou.' Somali *āḍ(i)* Saho *atu*, etc.; so also in some verbal prefixes and suffixes of the second person / Berber: Old Numidic and Modern Berber prefix of the second person *t-* / probably in Egypt. *ṭw ... tw* 'thou' (masc.), *ṭn ... tn* 'thou' (fem.), if one does not assume *t* < *ṭ* < *k* // Uralic: *\*tyna ~ \*tind* 'thou': Fin. *sind* 'thou' (*si-* < *\*ti-*), Mordvinian *ton*, Hung. *te*, etc. / Samoyed: Nganasan *tannan* 'thou', Selkup *tan*, *tat*, Enets *todi* Kamasin *than* 'thou', and everywhere the verbal ending of the second person. / Yukaghir *tet*, Chuvantzy *tota* 'thou' // Alt.: Mongolian *ci(n)-* < *\*tin* 'thou', *ta(n)-* 'you' // Chukchee-Kamchatkan: East Kamchadal (Kamchatka River dialect) *tue* 'thou', West Kamchadal *tuzanū* 'you', Chukchee *ṽa-t* 'thou', *turi* 'you', Koryak *ṽa-čča* 'thou', *tuū* 'you.'

S: Indo-European: verbal ending of the second pers. sg. *\*-si*. *\*-s* // Alt.: Turkic: Old Uighur *sañ*, etc. / Tungus: Evenki, Manchu *si* (obl. case *sin-*), Gold *si*, etc. // Chukchee-Kamchatkan: East Kamchadal *suze* 'you', *-z* in Itelmen *ki-z(a)*, *k-za* 'thou' (in all dialects) possibly from *\*z* or *\*d* // Kartvelian: Chan *si(n)* 'thou', probably also in Mingrelian and Chan *skan-* (possessive pronoun of the second pers.) and in Georgian *šen* 'thou' // Sumerian: *zae*, *za* 'thou', eme-SAL *zé* 'thou.'

K: Sem.-Ham.: Chadic, Hausa *kai*, *kē* 'thou', Logone *kān*, etc. / Berber: Ahaggar *kai*, *kem* 'thou', Shilha *kij(in)*, *kem(in)* 'thou' / Egypt. *k* — pronominal suffix of the second pers. sg. masc. / Sem. and Cushitic *k* in the possessive and object suffixes of the second person // Kartvelian: in the verbal object prefixes of the second person: Georgian *g-*, Chan *g-*, Mingrelian *g-* ~ *r-* Svan *ǰ-*.

### 4) 'WHO', 'WHAT'

K: Indo-European *\*k<sup>w</sup>i-*, *\*k<sup>w</sup>o-*, *\*k<sup>w</sup>-ā*. // Uralic *\*ku* 'who'<sup>14</sup> (Fin. *kuka*, Kazym Ostyak *həj*, etc. / Samoyed.: Selkup *kudō*, *kutte*, etc.), *\*ke ~ \*ki* 'who'<sup>15</sup> (Fin. *ken*, Mordv. *kue*, Hung. *ki*, etc. / Ykaghir *kin*). // Alt.: Turkic: Old Uighur *kim* 'what', etc. / Mong. *ken* 'who' / Tung.: Gold *haj* 'who', etc. // Chuk.-Kamch.: West Kamchadal *k'e*

'who', South Kamch. *kač* id., East Kamch. *ket* id. // We do not include *K* in several Cushitic languages (Kaffa *kōn* 'who', Somali *kuma* 'who', Galla *kan* 'which') and in some Chadic languages (Bura *ga* 'what', Somray *kana* id., Sokoro *kankema* id., Logone *γam*, *γwani* ~ *γoni* ~ *woni* id., etc.) due to territorial limitations.

M: Ham.-Sem.: Arab. *man* 'who', *ma* 'what', Hebr. *mā* 'who', *ma* 'what', Akkad. *mannu* 'who', *minū* 'what', etc./ Eg. *m* 'who, what'/Cush.: Sidamo *ma* 'what', Saho *mā*, *mā* 'who', etc./Berber: Tashelhit *ma* 'who, what', etc./Chad.: Hausa *mi*, *me* 'what', Angas *me* 'what', Buchuma *mino* 'what', etc. // Ural. *\*mi*:<sup>16</sup> FU 'what': Fin. *mikä*, Mordv. *meže*, Komi *mij*, Kazym Ostyak *muj*, Hung. *mi* / Samoyed: Nganasan *ma* 'what', Enets (Yenisei Samoyed) *mii*? 'what' // Chuk.-Kamch.: Chuk. *menin*, *mik-* 'who', Koryak *me:ki* 'who', West Kamch. *min* 'which', *menkanan*, *minein* 'what' (Adj.) // Kartv.: Chan and Mingrel. *mi(n)* 'who', *mu-* 'what', Svan *māj-* 'what' // The radical in *M* is also present in the IE languages (Hittite, Tocharian) and in Altaic languages (cf. Chuvash *mān* 'what').

#### 5) 'TONGUE'

KR: Ham.-Sem.: Chad.: Gabri *kelendin* 'my tongue', Nangire *kelendem* id. Kaba *klāndi*, Mojel *kelesūm* id., Gidder *hirdiko* 'tongue', etc. / Cf. Sem.: Arab. *qāla* '(he) said', Hebr. *qōl* 'voice' / Eg. *hrw* 'voice' // Ural. *\*kele*:<sup>17</sup> Fin. *kieli* 'tongue', Erzja *kel* Votyak *kil* id., etc. / Samoyed: Ene (Yenisei Samoyed) *sioro*, *sioḍo* 'tongue', Selkup *šel*, *sie* and so on. // Alt.: Mong. *kelen* 'tongue' / Tung. *\*hilnu*, (so Benzing 1956:991; Swadesh 1962a:295, 305) 'tongue' > Gold *sirmu*, *sinmu*, Manchu *ilenu* etc. / Cf. also Chuvash *kala* 'speak, say' and Old Uighur *kāl-māčī* 'interpreter' // Chuk.-Kamch. *\*gil*:<sup>17</sup> (so Swadesh 1962), 'tongue' > Chuk. and Koryak *-jil-/jel-*, Kamch. *-čil-*, *čal-* (West Kamch. *l'āčal*, *l'čil*, *ičil*, East Kamch. *dičel*, South Kamch. *ničil* - traces of a reduplication?) // Ural.: Yukaghir *azū* 'tongue', Chuvantzy *aīže* id. // Kartv.: Georg. *ena* 'tongue' // Sumer. *eme* 'tongue'

#### 6) 'NAME'

NM: IE: OI *nāman* 'name', Gr. *ῥόνομα*, Lat. *nōmen*, Toch. *ñom*, etc./ Ural. *\*nime*:<sup>18</sup> 'name': Fin. *nimi*, Zyrian *nim*, Vogul *nam*, etc./Yurak *num̃*, *nim̃*, Nganasan *ñim*, Kamassian *nim*, etc./ Yukaghir *niu* 'name'. // Cf. Sumer. *inim* 'word.'

N. . . : Chuk.-Kamch.: Chuk. and Koryak *nənn*- 'name.'

S. . . : Ham.-Sem.: Sem. \*šm 'name' : Arab. ʔism, Hebr. šēm, Ethiop. sām, etc. / Cush.: Omoto *sunča*, Walamo *sumta*, Sidamo *sumʔe* (provided this is not a šemitism), Bilin *sun*, Hamir *zun*, etc. // Kartv.: Georg. *saxeli*, Svan *žaxe* 'name.'

#### 7) 'EYE'

Ø. . . : IE: OI. *akši*, Lat. *oculus*, OCS *oko* 'eye', etc. // Ham.-Sem.: Sem. \*ʕjn 'eye' / Eg. ʕn, ʔr.t / Cushitic radicals of the type ØR and ØN: Bilin ʕil, Sidamo *illé*, Somali *il*, Mbugu *ila*, etc., Afar *inti* 'eye', Somali *enḏo* 'eyes', etc. / Chadic roots of the types ØT and ØR: Hausa *ido*, Jonkor ʔidi, Jegu ʔudé, Karbo *úden*, etc., Gulfei *el*, Dari ʔr, Muzgu *arai*, Matakam *ere*, etc.; cf. also Cush. Ø(J)P: Kaffa *afō* 'eye', Omoto *aufe*, *aype* 'eye', etc. // Ural.: Yukaghir *anze* 'eye' // Sumer.: Eme-KU *igi* 'eye', Eme-SAL *ibi* (or *ide*?) 'eye.'

#### 8) 'HEART'

KR: IE \**ḱerd-*: Hitt. *kard(i)*, Gr. *καρδία*, OCS *srŭdice*, etc. // Ham.-Sem.: Arab. *k'lb* 'heart' / Chad.: Mubi *kōrlq* 'heart', Gudu *quraksə*, etc. // Alt. (root with the form \**ǰ'ir-*): Mong. *ǰürüken* 'heart' / Turkic: Old Uighur *jüräk*; it seems that \**ǰ* here probably originates from \**d* (root TR), but it may as well have originated (as indicated by the general typology of phonetic changes) from \**q* (then the root is KR), or from \**j* (root of the type JR, - see note 11). // Chuk.-Kamch.: East Kamch. *gulḡ* 'heart' // Kartv.: Georg. *guli*, Chan and Mingrel *guri* 'heart'.

KT: Ural.: Fin. *sydän*, *sydäme-*, Skolt Lapp č̣e, Votyak *sülem*, West Cheremis *šüm*, Hung. *sziy*, etc. / Samoyed: Motor *kejem*, Taigi *kejm*, Selkup *šd*, Yurak *sej*, etc. B. Collinder (1960:409, 1955: 146) reconstructs Proto-Uralic \**šuδV* 'heart.'

RP: Ham.-Sem.: Sem. \**lhb* / Cush. Galla *luba/u*, Kwara *lābākā*, Bedaue *lēb*, etc. / Chad.: Klesen *réwu*, Makeri *irfū*, etc. / Probably Eg. *ih* 'heart.'

R. . . : Chuk.-Kamch.: West Kamch. *linḡyč* (after Dybowski), Chuk. and Koryak *linlin* 'heart.'

MP: Alt.: only in Tungusian languages: Gold *miavan*, Evenki *mēvan*, etc. It is unclear if the root is old; therefore it is not calculated.

#### 9) 'TOOTH'

TN: IE: OI *dant-*, Lat. *dent-*, Got. *tunþus*, etc. //

Ham.-Sem. SN or TN: Chadic languages show fricative or lateral  $\theta$ : Basa  $\theta eni$  'tooth', Ngala  $\theta \alpha \eta$  Logone  $\lambda ini$ ,  $\lambda \alpha n$ , Cabin  $\lambda en \bar{a}$  ( $\lambda$  = voiceless lateral fricative, similar to  $\theta$ ); in other languages,  $\xi$ ,  $s$ ,  $l$  correspond to our  $\theta$ ,  $\lambda$ : Klesem  $s \bar{a} ni$  'tooth', Banana  $s i a n u$ , Muzgu  $\xi \bar{e} \xi i e \eta \sim \xi i \eta$ , Semitic  $* \xi n$ , Holma  $l i n$ , Buchuma  $l i n t o$ , etc.

T. . .: Ural.: Samoyed: Kamassian  $t h i m \bar{a}$ , Motor and Koibal  $t y m e$ , Selkup  $t u m \sim t e n \sim \xi e n$ , Nganasan  $t i m i$ , Yurak  $t i b \bar{a}$ ; a root in T. . . is also present in Yukaghir  $t o d i$ , Chuvantzy  $t e d e g i$  // Alt.: Turkic: Old Uighur  $t i \check{s}$  'tooth', Turkish  $d i \check{s}$ , etc.

ST: Alt.: Mong.  $\xi i d \bar{u} \bar{n}$  'tooth' // Kartv.: Svan  $\xi d i k$  'tooth'

S. . .: Sumer.  $s u$  'tooth.'

KP: Kartv.: Georg.  $k' b i l i$ , Chan  $k' i b i r$  'tooth' // Chuk.-Kamch.: East Kamch.  $k i p t$ , South Kamch.  $k e p k e p$ , West Kamch.  $k i p k i p$ .

KS: Ham.-Sem.: Chad.: Dormo  $g a s e n e$ , Gabri  $k s o n g$  / Cush.: Beɛaue  $k \bar{o} s$ , Kaffa and Mocho  $g a \check{s} o$  / Probably Berber: Tashelhit  $a h \bar{u} s$  'tooth.'

10) 'NO(T)' (verbal negator and prohibitive)

ØR: Ham.-Sem.: Sem.: Hebr.  $\text{?} a l$  (prohib.), Ethiop.  $\text{?} a l$  (neg.), Akkad.  $\text{?} l$  (neg.) / Berber  $u r$  (prohib. and neg.) // Ural.  $* e l \bar{a}$  20: FU prohib, morpheme: Fin.  $\bar{a} l \bar{a}$ , Est.  $\bar{a} r a$ . North Lapp  $\bar{a} l \bar{a}$ , Ostyak  $\bar{a} l$  etc./Kamassian  $e l (e)$ - (root of the prohibitive verb) / Yukaghir  $e l$  (neg. and prohib.) // Alt.: Mong.  $\bar{u} l \bar{u} i$  (neg.) // Kartv.: Georg.  $a r$  (neg.).

Ø: In addition to the ØR-cases, the following morphemes are present: Ham.-Sem: Sem.: Ethiop.  $\text{?} i$  (neg. and prohib.), Akkad.  $a i$  (prohib.) / Chad.: Buchuma  $a$  // Ural.:  $* e$  21: root of the prohibitive verb in FU languages: Fin.  $e$ -, Erzja  $a$ -, Votyak  $u$ -, etc.; Vogul prohibitive particle  $a t$ , etc. / Samoyed: Nganasan prohibitive  $i$ -, Selkup neg.  $a s$ ,  $a s s a$ ,  $a \check{s} s a$  // Alt.: Neg. root  $* e$ -. Tung. prohibitive verb: Evenki  $e$ -, Lamut  $e$ -, Gold  $e$ -/  $a$ -, Ude  $\partial$ -, etc. / Mong.  $e s e$  / Probably also Turkic neg. suffix  $-a/e$ - // Chuk.-Kamch.: Chuk. and Koryak neg. prefix  $a$ -/  $e$ -, Chuk. neg. particle  $e t l e$ , Chuk. and Koryak  $u j \eta e$  (neg.), Chuk.  $e r \eta e$  (prohib.), etc.

N: IE  $* n e$ ,  $* n$ - // Ham.-Sem.: Eg.  $n$ , Coptic  $\partial n$  / Cush.: Somali  $-i n$ ,  $a n$ ; Ometo  $-e n n a$  / Chad.: Probably, Bodo  $n \bar{a} b e$ , Budduma  $n g \partial$  (prohib.) // Ural.: Hung.  $n e m$  'not' and  $n e$  (neg.), Samoyed neg. verb: Yurak  $n i$ -, Nganasan

*ñe-*, Enec (Yenisei Samoyed) *ñi-* // Kartv.: Georg. prohibitive *nu*, Svan prohib. *nom* // Sumer.: Neg. *nu*, prohib. *nu* and *na*.

M: IE neg. *\*mē*: OI *mā*, Avest. *ma*, Gr. *mē*, Arm. *mi*, Toch. neg. *mā* // Ham.-Sem.: Arab. *mā* 'not' / Eg. prohib. *m* / Cush.: Afar *m-*, etc. // Phonetically similar is Turkic prohibitive *ma/mā*,<sup>22</sup> but it is possible that it originated from a nominal suffix *m* + neg. *e-*, see Ramstedt (1924).

#### 11) 'FINGERNAIL'/'TOENAIL'

PR: Ham.-Sem.: Chad.: Mandra *feliṣe*, Jegu *illō*, Pika *pala*, Bura *m-pil*, Hausa *farče* (also 'toe') / Probably Sem.: Arab. *ṣfr*, Hebr. and Akkad. *špr*, Ethiop. *šfr*, Aramaic *ṭpr* / Cush.: Possibly in the meaning 'finger': Somali *far*, Afar *ferā*, and so on. // Kartv.: Georg. *prčxili*, Mingrel *bircxa* / Cf., in the meaning 'finger': IE: Toch. *prār*, Lith. *pirštas*, OCS *prŕstŭ* / Ural.: Nganasan *bāra btā* 'little finger' / Alt.: Manchu *ferhe* 'thumb', etc.

P. . . or W. . . : Chuk.-Kamch.: Chuk. and Koryak *-vaq-/væg-* where *-v-* originates either from *\*p* or from *\*w*.

K + nasal (most probably KM; often KMS/KMT): Ham.-Sem.: Cush.: Galla *k(i)ensa*, Kaffa *gāmō*, etc. / Chad.: Hausa *k'umba*; also possibly Budduma *kuddi* (with the loss of internal nasal) // Ural.: FU *\*kŭnče ~ \*kŭče*: Fin. *kynsi* (stem *kynte-*), Moksha-Mordvin. *kenže*, Ostyak (Konda and Yugan dialects) *konč*; with a loss of nasal: Cheremis *kāč*, Votyak *qīžī*, etc. / Samoyed. (with a loss of nasal): Kamassian *kāda*, Selkup *kač*, *kat*, *kače*, Nganasan *katu*, Enec (Yenisei Samoyed) *koda*, Yurak *hada* // Alt.: Mong. *qumusun* / Also probably (with a loss of nasal): Gold *hosakta*, Orochi *hosikta*, and other Tung.

KK: Chuk.-Kamch.: West Kamch. *kichun*, *kuchkuch* (Dybcwski).

#### 12) 'LOUSE'

WT: IE: Lith. *utė*, Latv. *uts*; a similar root of the type WS is present in OCS *vŕši*, Lith. *vievesa* and probably in OHG and OE *lūs* // Ural. (root WNT or WMT): Selkup *unž*, *undže* (Castrén), Enec (Yenisei Samoyed) *addu*, Kamas. *ūnū*, Nganasan *nomt(t)un* (< *\*unt-* or *\*umt-*).

Similar roots mean 'nit' in Ham.-Sem. (berber *iutt*), Chuk.-Kamch. (South Kamch. *vudud*). -Cf. Turkic *bit* 'louse' which, according to our automatic notation, should be

written as PT. Cf. Cush.: Kwara *betā* Hamir *bettā* 'louse.' In Lamut *ut* 'nit', suffix *-t* is either old and original or results from a re-etymologization).

Ø + nasal (ØM, ØMT, ØNT, ØN): Ham.-Sem.: Chad.: Jegu *ʔintāatō*, Banana *āndārā*; probably also Logone *mtāsə* and Mubi *idēdi* / Cush.: Galla *inʒiran*, Somali *inʒir-*, and so on. / Semitic root *\*ʔnb*: Akkad. *nābu* 'louse', Hebr. *ʔinəbbā* 'nit', etc. // In IE, a similar root designates 'nit': Arm. *anic*.

KM: Ham.-Sem.: Arab. *k'ml*, Ethiop. *k'wml* / Eg.: Saidic Coptic *komf*, *koomef* / Cf. Hausa *k'uma* 'flea' // Alt.: Evenki and Lamut *kumke*, Orochi *kumuha*, Ude *kumugə* // Cf. IE root of the type KNT (*\*qonid*, *\*gnid*, *\*(s)qnid*) with the meaning 'nit': Gr. *kōnis*, *konidos*, OE *hnitu*, OCS *gnida*, etc. - Cf. Somali *gindēʔelki* 'small louse'. - A root KN is present in Chuk.-Kamch. languages as well: West Kamch. *chynym* 'nit' (according to Dybovski).

TR: Ham.-Sem.: Bohairan Coptic *tersis* // Kartv.: Georg. *t'ili* 'louse', Chan *t'i*, etc.

T(R): Ural.: Root TR appears in the meaning 'nit': Yurak *talik*, Kamas. *thāri*, etc.; in the meaning 'louse' a root *\*tāje* appears in Finno-Ugric: Fin. *tāi*, West Cheremis *tij*, Votyak *tej*, Hung. *tetü(v)*, Kamas. *thāri*, etc. [but cf. above]. Of course, in this set of data only the initial T is taken in consideration. // Alt.: Evenki *tile-* 'louse', Gold *čile-* 'look for lice', Olchi *tiktē-* 'louse', Gold *čiktē* id., etc.

13) 'TEAR' (noun): No correspondences.

14) 'WATER':

W(N)T: IE: Got *watō*, OCS *voda*, Hitt. *watar* etc.; *\*wend-* in: Lith. *vanduo*, and so on. // Ural. *\*wete*: Mordv. *veđ*, Cheremis *vəđ*, Hung. *viz*, etc. / Samoyed: Forest Yurak *wit*, Selkup *ūt*, *öt*, etc. / Yukaghir: Chuvan. *onde*, Yukaghir *ōzi* (according to Jochelson), *oʔi* (according to O. Tailleux).

W: Alt.: Mong. *usun* 'water' / Cf. Evenki *ūakma* 'water', Gold *uama* 'wave.'

MR: Alt.: Evenki and Lamut *mū*, Gold *muə*, Manchu *mu-ke* 'water', etc. / Cf. Korean *mul* 'water' / Cf. Mong. *mören* 'river'<sup>23</sup> // Chuk.-Kamch.: Chuk. and Koryak *miməl* 'water' (redupl.), West Kamch. *meml-* (according to Jochelson). // Cf. IE *\*mōri* 'sea.'

M: Ham.-Sem.: Sem.: Arab. *māʔu*, Hebr. *mayim*,

Ethiop. *māy*, Akkad. *mū*, etc. [s. above]. / Eg. *mw*, *mhj*<sup>24</sup> / Chad.: Lame *mhi*, Banana *mbona*, Malabu *mbaiye*, etc. It is possible that Chadic and Berber morphemes of the type *ØM* are related (s. below).

*Ø* . .: Sumer. *a*, *eš* 'water' // Chuk.-Kamch.: East Kamch. *azamch*, *ažimch*, West Kamch. *i* (after Jochelson), South Kamch. *ii*. // Ham.-Sem. (root of the type *ØM*): Berber: Tashelhit *aman* / Chad.: Angas *am*, Fika *ama*, Karekare *amu*, Bode *amu*, Ngodjin *am*, Logone *am*, Barein *ami*, etc. Probably, morphemes of the type *JM* belong here as well: Muzgu *yem* 'water', Mutura *iam* id.; they are comparable with Sem. *jm* 'sea', Yurak *jam* id. and Nganasan *jam* id.

#### 15) 'DEAD'

*M* . .: Ham.-Sem. MT: Sem. *\*mwt* / Eg. *mt* / Chad.: Hausa *matačče*, Jegu *moot* (in the meaning 'die': Logone *mti*, Gulfei *made*, Doai *emči*, etc.). / Berber: Tashelhit *ěmmet*, etc. (provided this is not an Arabism). Beside the above roots of the type *MT* there are roots with initial *M*- but differing as far as the second consonant is concerned: Eg. *mʷw* 'dead', *mnj* 'die'. Chad. (type *MR*): Somrai *márfälin* 'dead', Muzgu *amra* 'death', *mara* 'die', etc.<sup>25</sup>

*MR*: IE: OI *mrtu-*, Lat. *mortuus*, OCS *mrŭtvŭ*, etc.

*KR*: Ural.: Fin. *kuollut*, Erzja *kulož*, Votyak *kuləm*, Hung. *halott*, etc. / Samoyed: Yurak dial. *halmer*, etc. ( < Ural. *\*kole-* 'die'). // Kartv.: Svan *lūdāār* 'dead', cf. Mingrel - *γur-* 'die', Chan-*γur-* id., Svan *dagra* 'die' // ?Alt.: Turkic: Altai-Kiži *kal-* 'die' (it is not clear if the meaning is original; this example is not taken in consideration in the statistics).

*K* . .: Chuk.-Kamch.: West Kamch. *qisaknen* 'dead', etc.

*WR* or *ØR*: Alt.: Turkic: Chuvash *vil(n)ě* 'dead', Old Uighur *ölüa*, etc. / Cf. Mong. *ölbüri* 'epidemy, plague' / Cf. Evenki *ülbü-mi* 'I am dying of hunger.'

*W* . . or *P* . .: Chuk.-Kamch.: Chuk. *wʷilən* 'dead', Koryak *wiʷelʷən* id.

*W* . .: Sumer. *uq*, *uš* 'dead.'

*P* . .: Alt.: Mong. derivative from *ükü* 'die < *\*püh-ü*, as in Monguor (Tungshiang) *fugu* 'die.' // *P* is the initial in a Tung. root presented by Evenki *bučē* 'dead', Gold *bujkin*, Orochi *bukine*, Manchu *bučehe*, etc.

It seems certain that not all phonetic correspondences

will prove to be etymologically correct cognate sets. This is not too important for us now, but what is important is the fact that the amount of correspondences so far observed by far surpasses the possible amount of chance correspondences; this means, of course, that among the correspondences here there are those in which similarity cannot be explained by chance.

Using a mathematical formula one can calculate the probability of a chance similarity that would be not less than the amount of correspondences actually observed with a given meaning. Table 1 shows that tests applied to 13 out of the 15 stablest meanings result in figures not less than 10-2; this leads to a figure 10-20, which characterizes the probability of chance coincidence for our data in general. This proves, beyond any reasonable doubt, that the observed similarity cannot be explained by chance.

Is it possible to explain these similarities by borrowing? Hardly, since most of the meanings cited above contain words that are seldom borrowed. Our results show that in 200 languages of Europe, Asia, and Africa,<sup>26</sup> words that have been analyzed here with the meaning 'I/'we', 'thou/'you', 'tongue', 'eye', 'no(t)', 'louse', 'tear (noun)', 'water' never undergo borrowing. As for the value 'tooth', only two cases of borrowing have been observed, namely Monguor (Tunghsiang) from Chinese and, possibly, Baltic-Finnish from some Indo-European language. There are only three cases of borrowings in the word group with the meaning 'fingernail'/'toenail' (Caxur from Azerbaidzhani; Malayalam and Tamil from Sanskrit); not more than four cases of borrowings in the set for 'dead' (Japanese from Chinese; Somali and, possibly, Berber from Arabic; in Kannada, one of the words with this meaning has been borrowed from Sanskrit); same in the meaning set for 'two' (in Japanese, Korean, Monguor and Berber—borrowings from Chinese and Arabic, accordingly, co-exist with original words with the same meaning). Exceptions are words for 'heart' and 'name': there are nine to twelve borrowings in the first case (Japanese, Vietnamese, and, possibly, Korean from Chinese; Hindi and Sindhi from Persian; Dravidian languages from Sanskrit); and about ten borrowings in the second case (Arabic into some Turkic, Iranian, and, possibly, Berber languages; one of Akkadian synonyms

- from Sumerian).

Now, let us assume that in a given meaning for which two hundred languages show no borrowings, a borrowing can be found in one of an additional hundred languages. In this case, a set of borrowings capable of explaining coincidences in Sibero-European languages has a probability of less than  $10^{-14}$ . Even if part of the similarities are said to be borrowings and another part are said to be caused by chance, any combination of borrowings and fortuitous coincidences capable of explaining the above similarities of Sibero-European languages will have an infinitely low probability.

All this leads us to conclude that the correspondences between different Sibero-European languages can be explained neither by chance nor by borrowing. The only possible explanation is the existence of a genetic relationship between (at least, several) families of Sibero-European languages. In order to find out which of the Sibero-European language families are mutually related, one can undertake a series of binary comparisons (on the basis of the above list of semantic values): one can compare any two families, or a certain family with a certain number of families, combining such a comparison with calculations similar to the calculations mentioned above. In doing so, one can arrive at the following figure: a probability of fortuitous coincidence (which is not higher than the actual coincidence) of Indo-European and Uralic morphemes is less than 0.002; that of Hamito-Semitic morphemes with Indo-European and Uralic morphemes is less than 0.005; that of Altaic and Uralic morphemes is less than 0.001; that of Chukchee-Kamchatka and Uralic morphemes is less than 0.015; that of Kartvelian morphemes with Indo-European, Uralic and Hamito-Semitic morphemes is less than 0.002. As for Sumerian, the upper limit of fortuitous coincidence of Sumerian morphemes with Uralic ones is 0.06 (or even 0.17, provided the EME-KU reading  $\eta a, \eta ae$  'I' is correct and that the morpheme  $\eta-$  here is more archaic than  $m-$  in EME-SAL *me* 'I').

Thus, we can speak with more confidence about a genetic relationship between Indo-European, Hamito-Semitic, Uralic, Altaic, Chukchee-Kumchatkan, and Kartvelian languages,<sup>27</sup> but with less confidence about a genetic relationship between these languages, on the

one hand, and Sumerian, on the other hand.

## NOTES

### Dolgopolsky Article

1. See Pedersen (1903), and also his report to the Third International Congress of Linguists (1935:328-33)
2. The term "Nostratic," from Lat. *noster* 'ours', is evidently unfortunate.
3. A specific source of similarities in different languages is the so-called "*affinité élémentaire*", that is, a purely accidental coincidence, but one which results from conditions that make such coincidences highly probable, cf. onomatopoeic terms, interjections, and the common lexicon of *Kindersprache*. We do not consider words from these sectors in our comparisons.
4. Here, we will not take up the problem of a possible relationship between Kartvelian and the Northern Caucasian languages. Then, too, we do not employ the Northern Caucasian languages when evaluating the Sibero-European hypothesis, since their relationship (provided it is genetic) with both Kartvelian and Indo-European is overshadowed by further developments. Nevertheless, it is possible to find Northern Caucasian radicals that are similar to those in Indo-European, cf. a root of the type *MS*, both 'moon' and 'sun': Adyghan *maze* 'moon', Abkhaz. *ā-mza* id., *ā-mra* 'sun'; Botlix. *mihi* 'sun', etc. In Kartvelian, cf. the following: Megrel *mze* 'sun', Svan *məž* id. // Semitic \**šamš* 'sun', Chadic: Bata *mōsē* 'sun', etc. // Eskim. (Naukan) *masak* 'sun' // IE \**mēs*, \**mēnes*- 'moon.' For a radical of the type *NT* 'louse', note: Lezgin *net*, Lak. Avar *nac* id., etc., cf. IE: OCS *gnida*, Gk. *konid*- 'nit', etc. // Uralic: Selkup *undže* 'louse', etc. // Hamito-Semitic: Galla *enžīži* 'nit', *anžiran* 'louse', Chadic: Iegu *ʔintáátó* 'louse', etc. Note the prohibitive morpheme *M*: Lak *ma*, Dargh, *-ma-*, Adygh. *mə-*, etc. Finally, cf. Kartvelian, Indo-European, and Hamito-Semitic negative particles of the type *M* as given below.

The Yukaghiran language, together with Chuvan and Omok, is considered a member of the Uralic family; at least this is the conventional wisdom. It is a well known

fact that various scholars, e.g., I. Ankeria and E. A. Kreinovič, seriously opposed B. Collinder, O. TAILLEUR, and K. Bouda, who had assigned Yukaghiran to Uralic. I would prefer to treat Yukaghiran as a separate family. The scheme proposed here would lose nothing by this: we would then consider just eight and not seven language families, and the chance of a purely fortuitous coincidence would then be even lower. As for Korean, we do not consider it here. Eskimo-Aleut languages are also omitted, and it appears that they belong to Sibero-European.

5. The essence of Lees' work is presented in Swadesh (1952:452-63; 1955:121-37).

6. In doubtful cases, a semantic value is considered to have 0.5 replacements.

7. Data for this meaning are absent in several languages.

8. According to a preliminary statistical assessment, the consonants *c*, *č*, *z*, *ž* usually originate from *\*k* and *\*g*; less frequently *c*, *č* originate from *\*t*; *ž* from *\*j*; and *z*, *ž* from *\*d*.

9. Here and subsequently, both Proto-Uralic and Proto-Finno-Ugric reconstructions are cited after Collinder (1960:405-15).

10(E). For Uralic and Altaic correspondences, cf. Illič-Svityč (1976:63-4); all are from Nostr. *\*mi* 'I,' oblique from *\*mi-nV*.

11. Most scholars give this form. If one follows R. Jestin and A. Falkenstein, one can assume *na-e*, *na* 'I' (EME-KU), *me* 'I' (EME-SAL), and this lowers the probability that Sumerian belongs to Sibero-European, provided, of course, EME-KU is the more archaic of the two forms.

12(E). This has also been reconstructed by Illič-Svityč (1976:66) as *\*mi*, now see Dolgopolsky (forthcoming). In accordance with the author's introduction to this paper, the Hamito-Semitic forms do not belong here.

13. On phonetic reconstructions of Chukchee-Kamchatkan, see Swadesh (1962). This thesis about the nasal prefix was formulated by Swadesh in an unpublished paper, "American-Eurasian Linguistic Relations" (1960).

14(E). Cf. Illič-Svityč (1971:355-6): Uralic *\*k<sup>o</sup>/u-* 'who', IE *\*k<sup>w</sup><sub>o</sub>-* *íd.*, Hamito-Semitic *\*k(w)/q(w)* *id.*, etc., all from Nostr. *\*K'o* 'who'.

15(E). Cf. Illič-Svityč (1971:348-9): Uralic *\*ke-* 'who'. Illič-Svityč believed that this Uralic/Altaic pronoun with

\*e indicated proximal objects and was thus in opposition to \*K'o which indicated distant objects, cf. Uralic/Altaic \*e- 'this' vs. \*o-/a- 'that', see Illič-Svityč (1971:258, 271-2).

16(E). Cf. now Illič-Svityč (1976:66-68) where, *inter alia*, we find: Nostr. \*mi 'what' > Uralic \*mi id., Turkic \*mi id., Hamito-Semitic \*m(j) 'what, who', Kartv. \*maj 'who', IE \*mo-, the stem of interrogative adverbs.

17(E). Cf. Illič-Svityč (1971:221-2) where, *inter alia*, we find: Uralic \*kēle, Turkic k'älä-, Mong. \*kele(n) < Nostr. \*K'ülHä 'tongue.'

18(E). Cf. Illič-Svityč (1976:82-3) where, *inter alia*, we find: Uralic \*nime, IE \*nōm-, Hamito-Semitic (?) \*nb- (if dissimilated from \*nm-) < Nostr. \*nimi 'name.'

19(E). Cf. Dolgopolsky's remarks in the introduction to this paper.

20(E). Cf. Illič-Svityč (1971:263-4) where, *inter alia*, we find: Uralic \*āla/ela (2nd sg. imper. of neg. verb), Drav. \*al(a)- (neg. verb), Mong. (?) \*ūle (negation), Hamito-Semitic \*ʔl/\*lʔ (prohibitive and negative particle). (?) IE: Hitt. lē (prohibitive particle), (?) Kartv.: Georgian ar(a) 'not', all from Nostr. \*ʔāiā, a particle signifying categorical negation.

21(E). Cf. Illič-Svityč (1971:264-5) where, *inter alia*, we find: Uralic \*e- (negating verb), Drav. \*-a-/e- (infix of negative verbs, negating infix), Mong. \*e-se- (stem of negative verb), Tung. \*ā- (negative verb), (?) Turkic: Chuvash an (< \*e-n, negative particle), Hamito-Semitic \*ʔj (negative and prohibitive particle), all from Nostr. \*ʔe (negative particle).

22(E). Cf. Illič-Svityč (1976:56-7) where, *inter alia*, we find: Turkic \*-mā- (negative infix in verbs), Drav.: Gondi ma/inni (prohibitive), etc., IE \*mē (prohibitive particle used with imperative), Hamito-Semitic \*m(j) (prohibitive and negative particle), Kartv. \*mā/ō (prohibitive and negative particle), all from Nostr. \*mā, prohibitive particle.

23(E). Cf. Illič-Svityč (1976:60-1) where, *inter alia*, we find: Mong. \*mōre(n) 'river', Middle Korean mŭl, Korean mul, mur, mir 'water', Drav.: South-Dravidian \*mar-ai- 'rain' (borrowed by Sanskrit as māri- 'rain'), IE \*mor- 'bog, body of water', Hamito-Semitic \*mr 'moist/wet, rain, body of water', Kartv. \*mar-(ei) 'lake, moist soil, cloud', all from Nostr. \*māra '(having) liquid.'

24(E). Cf. Illič-Svityč (1976:62-3) where, *inter alia*, we

find: Hamito-Semitic *\*mw* 'water, liquid', IE *\*meu-* 'moist, moisten', Tung. *\*mδ* 'water', all from Nostr. *\*mEwV* (i.e., *\*mewV* or *\*māwV*) 'water, liquid'.

25(E). As far as the *r*-ful forms are concerned, cf. Illič-Svityč (1976:59-60) where, *inter alia*, we find: Nostr. *\*m(ā)rV* 'ill, be ill, die' > Hamito-Semitic *\*mr-* 'be ill', IE *\*mer-* 'die', Uralic *\*m(e)rV* 'wound, pain', (?) Middle Mongolian *mer* 'wound', cf. the Kalmyk interjection *mer-mer*, 'an expression of pain.'

26. Only loans from closely related languages have been taken into consideration. Each meaning contains but one word, the principal synonym.

27(E). Dolgopolsky did not employ Dravidian languages, and these were used by Illič-Svityč. In fact, as we have seen, East Nostratic is comprised of Uralic, Dravidian, and the Altaic languages. As for Chukchee-Kamchatkan, S. Nikolaev considers them to belong to a different macro-family, namely, "Macro-Dene-Caucasian", along with Yenisseian, Sino-Tibetan, North-Caucasian, Athapascan, Salishan, Wakashan and Algonquian languages. On the other hand, Eskimo-Aleut languages seemingly belong to Nostratic, see O. A. Mudrak, *K voprosu o vnesnix sv'az'ax ėskimosskix jazykov*. In: *Linguističeskaja rekonstrukcija i dreunejšaja istorija Vostoka. Tezisy i doklady konferencii. Čast' I*, (Moscow: Nauka, 1984), pp. 64-70. Cf. also the fifth volume of these materials: *Problemy izučeniya nostratičeskoj makrosem'ji jazykov*, (Moscow: Nauka, 1984) with papers by V. Dybo - V. Terent'jev, V. Ivanov, K. Pozdn'akov, and E. Xelimskij.

## IVANOV REVIEW I

V.V. Ivanov. Review of: V. M. Illič-Svityč, *Opyt sravnenija nostratičeskix jazykov (semitoxamitskij, kartvel'skij, indoevropej skij, uralskij, dravidijskij, altajskij). Uvedenie. Sravnitel' nyj slovar' (b - K')*. Moscow: In: *Ėtimologija* 1972. Moscow, 1974, pp. 182-4.

The main distinction between Illič-Svityč's work and other earlier or even contemporaneous attempts at a broader comparison of the principal language families of the Old World lies in the exceptional precision of his methodology. This can be seen from his scrupulous selection of material (even unclear and potentially dubious cases, particularly those connected with the descriptive, onomatopoetic character of corresponding lexical items and their isolated position within a given language family, have been carefully underscored by the author) which is ranged in well-conceived systems of phonological correspondences (for which examples are often provided in the dictionary from such thoroughly investigated languages as Slavic) and from the exhaustiveness of his preliminary investigations of data from within each of the language families that are compared. The late author of this dictionary informed me that the very restrictive nature of comparisons for the six language families was also determined by the same exigency of precision. Illič-Svityč did not exclude the possibility that some further languages (such as Hurrian and Urartian<sup>1</sup>—at least by virtue of Holger Pedersen's earlier hypothesis) might well belong to the Nostratic macro-family. Despite this realization, as a first step he considered it very important to supply theoretical foundations for the Nostratic hypothesis on the material basis of the most secure etymologies from these six families, the six families for which we have the most elaborate historical-comparative grammars. The extent of Illič-Svityč's insights into the details of each of these grammars is quite exceptional, so much so that the specialists in each of these families can find many new data in the entries of the dictionary.

Ilič-Svityč's work demonstrates the explanatory power of Nostratic theory by showing that a large number of

facts which remained inexplicable within the framework of a given language family can, in fact, be explained from the larger Nostratic perspective. For example, if one were to operate with Indo-European alone, then one could not provide etymologies for such words as Gk.  $\theta\alpha\lambda\alpha\varsigma\alpha$  and its Macedonian correspondents (cf. p. 216), Lith.  $\text{d}\check{\text{a}}\text{žias}$  (p. 219),  $\text{anq}\check{\text{a}}$  (p. 245), Latv. (<Kursham)  $\text{d}\check{\text{i}}\check{\text{n}}\text{q}\text{t}$  (p. 223), Sl.  $\text{da}$  (p. 215), and so on, which can be explained in the light of Nostratic theory. Especially interesting are those cases in which the semantic history of well-known subsets of comparable vocabulary are explained from the new vantage point conferred by Nostratic correspondences. One of the better examples might be Sl.  $\text{berq}$  whose meaning 'to take' happens to be an archaism (p. 176) and not, as was generally assumed, an innovation. Nostratic theory demonstrates that such semantic values as 'spring, young plants, young' for Sl.  $\text{jar-}$  are archaic (cf. correspondences in other Nostratic languages) whereas the value 'year' (as in many Indo-European languages) is an innovation (pp. 37 and IV). From among the various details of Indo-European correspondences that corroborate Nostratic reconstructions, the following fact is interesting. The original meaning 'wolf' for Nostr.  $*K\ddot{u}jna$  (p. 361), as reconstructed by Illič-Svityč on the basis of Uralic and Semito-Hamitic, is represented in Old Irish by  $\text{cú}$  'hound, wolf' (in the Irish sagas this word can designate a monster which kills cattle as well as people and threatens the  $\text{ulads}$ ), see Schmidt (1957:186). With respect to the Altaic languages (as well as the Japanese borrowing  $\text{inu}$ ,<sup>2</sup> cf. Tungusian forms of the type  $*\text{xina} > \text{ina}$ ), only the derivative, secondary meaning 'dog' (< 'domesticated wolf') is represented.

Particularly convincing are those results of external comparison which compel us to reconsider traditional Indo-European reconstructions which have already been cast in doubt by data from such Indo-European languages as Anatolian and Tocharian. In particular, the so-called "Brugmannesque spirants" belong here. Recently, Indo-Europeanists have been inclined to abandon them and to accept, on the basis of Anatolian and Tocharian data, a metathesized cluster comprised of alveolar + guttural (> guttural + alveolar, as in Gk.  $\chi\theta\omega\nu$  Skt.  $\text{ksam-}$ ; Hitt.  $\text{degan} < *che\hat{\text{g}}\hat{\text{o}}m$  'earth'). This hypothesis finds

brilliant confirmation in Nostratic etymologies for the words for 'earth' (p. 220)<sup>3</sup> and 'fish' (p. 219).<sup>4</sup> These examples are excellent illustrations of the secondary character of Indo-European stems affected by reshaping of the original vocalism. It is possible that a radical reconsideration of the theories of primordial vocalism will cause serious perceptual difficulties for scholars accustomed to traditional views, but it is precisely here that Nostratic comparisons will enable us to solve those very problems that have recently become so numerous in Indo-European studies.

Another corroboration that illustrates the relevance of Nostratic reconstructions for Indo-European linguistics is, for instance, the new interpretation of IE *\*sp-*, *\*st-*, *\*sk-*; as for *\*st-*, it explains, in particular, non-lenited Celtic *s-* in the group of words corresponding to Indo-European lexemes for 'love/guard' and to the Kartvelian set meaning 'need/sorrow.' For the semantics in Kartvelian, cf. Hitt. *ištark-* be ill with *št-* < IE *\*st-* < Nostr. *\*č-*, cf. p. 210. The fact that, at the Nostratic level, these clusters correspond to unitary consonants shows that their function as simple phonemes in Germanic poetry is probably very archaic (this would be in line with the exceedingly archaic character of Germanic consonantism which is relatively close to that of Nostratic, cf. e.g., the correlations between stops).

Illič-Svityč's dictionary enables the Indo-Europeanist (as well as specialists in the other five language areas employed by Illič-Svityč) to search for new facts in languages that were considered to have been thoroughly investigated. Suddenly, certain phenomena move to the foreground, phenomena which, when viewed within the framework of a given family alone, seemed rather trivial. There are many examples of such phenomena, but we shall cite but one of them from the sector of grammatical relations. Illič-Svityč discusses (p. 206) the format *\*-či-* in frequentative and iterative verbs. Its use in verbal suffixes with the significance "polypersonality" in Altaic (Literary Mongolian *-čaga-*/*-čege-*) undoubtedly requires comparison with Hitt. *-šk-* in its capacity as an affix denoting polyobjectivity. This is important for, among other things, an understanding of the typology of the semantic development of this formant.

An important contribution to the historical study of Indo-European syntax is Illič-Svityč's interpretation of extraordinarily archaic particles (one of which later evolved into the augment) as being simultaneously proclitic and enclitic (p. 251). This interpretation became clear for Slavists after V. Dybo's detailed analysis of Vasil' ev-Dolobko's Law that had anticipated probable comprehension of proclisis/enclisis in the Indo-European protolanguage, an interpretation accepted by other linguists after Illič-Svityč's death, see Jucquois (1970:540). This, as well as numerous other penetrating insights about the grammars of the language families considered and insights about their common Nostratic source are presented in the various entries of the dictionary in Illič-Svityč's typically laconic style.

Upon reading Illič-Svityč's work, one is left with a sensation—a sensation often evoked by a reading of any seminal scientific theory—of exceptional aesthetic perfection, a perfection achieved by having discovered certain principal correlations fundamentally in tune with the objects observed. Illič-Svityč himself clearly understood the obstacles in his way. The possibility that there were alternative explanations for some of the similarities he discovered—that might result from borrowing (such an explanation for some Indo-European/Kartvelian correspondences was proposed in recent studies by T. Gamkrelidze)—prompted Illič-Svityč to attentive investigation of some late contacts between the protolanguages he compared. A typical example was his article on Semitic loans in Indo-European. From among other observations of this kind, one can cite the problematic treatment of Svan *kārāxs* 'horn' as a possible loan from Indo-European (p. 351; and this interpretation also seems possible for Uralic *sarwo* 'horn', pp. XI, XXXI-XXXII, especially if one operates with a source in line with the *satəm* - type, p. 317), as well as his fine remarks about the migratory character of Near Eastern terms connected with construction techniques, e.g., Hitt. *kutt-* 'wall' (p. 317), cf. the clearly migratory character of Hitt. *pir*, *pam-* 'house', etc.

Those scholars who would follow Illič-Svityč in his Nostratic studies will benefit largely from those sections of his book in which he, with his customary scholarly

integrity, formulates issues which remained open for him. One such issue is the unclear correspondence between Uralic \*-k- (e.g., in Finn. *teke-*) and IE \*-h- (as in *\*dheh-* 'put', p. 224, cf. the comments about Uralic *\*jēγē*, p. 267) which can, it seems, be considered as support for the presence of other similar traces of Nostratic laryngeals in Uralic, see Rosenkranz (1966:173-5), Bańczerowski (1972). Particularly interesting here is the archaic nature of the conjugation of Finn. *tehdä* 'do, make', where, as in *nähdä* 'see', the -k- element appears in the 3rd pers. sg. and pl. (*hän teki*, *he tekivät*) and is similar to the reflexes of a laryngeal in pre-affixal position in forms such as Lat. *fēcī*.

It is worthwhile pointing out that, *en passant* in the dictionary entries, a whole series of brief comments are given for additional sets of words which are semantically similar to those that are discussed. These comments, albeit laconic, will be of tremendous value for future scholars. Thus, even in passing, Illič-Svityč remarks on the archaic semantics of Kartv. *ser-* (p. 241). His conclusion there can now be supported by "external" comparison with Hurrian *Šerri* (the name of one of the two bulls of the storm god, labelled 'Evening' and 'Morning', respectively; Chuvashan *šēr* 'night' seems phonetically different, lest one suggest borrowing from another language).

These disparate examples are cited only by way of illustration to reveal the perspectives which Illič-Svityč saw in his work, a work that outstripped the development of linguistics for decades to come. In a brief review one cannot even list all the particulars of this book, one of the highest achievements, not only of linguistics, but also of the whole corpus of twentieth-century humanitarian scholarship. It remains for me to express the hope that the next three volumes of the dictionary will be prepared with the same accuracy and conscientiousness that are worthy of their late author as was here displayed by V. A. Dybo, the volume's editor (who has supplied a very valuable introduction to the work as well as important tables of phonetic correspondences), and by Illič-Svityč's friends and colleagues who helped with the editing. All of those who will, little by little, become interested in this new field of linguistics founded by Illič-Svityč will impatiently await the publication of the next volumes.

## NOTES

## Ivanov Review I

1(E). As S. Starostin has recently demonstrated, both Hurrian and Urartian belonged to a different macrofamily: namely; Sino-Caucasian, or, broader still, Dene-Caucasian. For a definition of this latter term, see Foreword, pp. xvi. In a recent work, Ivanov (1983:155-7) has discussed a possible genetic unity for "North Caucasian," Hurrian-Urartian, Hattic, Yenisseyan, Sino-Tibetan, Birmanese, and, in North America, Na-Dene, as well as, possibly, some other language groups.

2. Polivanov (1960:179-80). A genetic relationship for Japanese and Tungusian forms is, on phonological grounds, out of the question.

3(E). According to Illič-Svityč both IE *\*dhq̥h-em-* 'earth' (with *-e-* in Balto-Slavic, Toch. B, and Albanian) and Kartv. *\*diqa* 'earth, clay' (Klimov reconstructs *\*tiqa*, but *d-* in Chan is original, whereas *t-* in Georgian results from assimilation with voiceless *\*q*) originate from Nostr. *\*diqV* 'earth'. IE *\*-ǵh-* is here due to assimilation after voiced *\*dh-*, namely, *\*dhq̥h-* < *\*\*dǵh-* < *\*diq-*; *\*g* was palatalized prior to the loss of *\*\*i*.

4(E). According to Illič-Svityč, both IE *\*dhǵh-u-H* 'fish' and Alt. *[d]iga-* id. (> Mong. *\*ǰiqa-[sun]* id.) originate from Nostr. *\*diqa* 'fish'. In Indo-European, the original stem, *\*\*diqa*, became *\*dhǵh-*, where palatal *\*ǵh* reveals loss of *\*\*i* (as in the word for 'earth'). Nostratic voiced stops regularly became Indo-European voiced aspirated stops.

5. For the above mentioned Gk. θάλασσα, see Levy (1972:295-300).

IVANOV REVIEW II

**V. V. Ivanov. Review of: V. M. Illič-Svityč, *Oput sravneni-ja nostraticeskix jazykov (semitoxamitskij, kartvel' skij, indoevropskij, ural'skij, dravidijskij, altaiskij). Sravnitel' nyj slovar' (1-3). Ukazateli, Moscow: Nauka 1976. In: Ėtimologija 1977. Moscow, Nauka 1979, pp. 179-84.***

The second volume of Illič-Svityč's Nostratic dictionary contains approximately a hundred entries (L-N) from the fully completed portion of the dictionary, as well as several other entries (P-3) from the initial, rough-draft version of his work. This means that all the entries prepared by the author himself before his tragic death have now been published; further entries will have to be extracted from Illič-Svityč's files. This will require much more extensive editorial effort. It is sad that the exceedingly difficult task of reconstructing the Nostratic lexicon has now been made even more difficult, for now a reconstruction of the author's intentions will also be required. This intentional perspective is in part reflected in both of the published volumes (and the editors made every effort possible to lessen this burden on the part of the reader).

During the five-year period between the appearance of the first and second volumes of the dictionary, several positive reviews have appeared, among which there are several by scholars who have themselves founded branches of comparative linguistics subsumed by Nostratics (e.g., B. Collinder, N. Poppe). As for the negative reviews, they are ill-conceived and, for the most part, reflect what is a typically emotional response to what is a fundamentally new discovery. Most of the original research published during these five years has been occupied with the binary comparison of language families, such as Indo-European vs. Semitic (A. Bomhard), Indo-European vs. Uralic (B. Čop). Such research holds a certain value even for Nostratics in general. As for the simultaneous comparison of a significant number of language families, with the exception of a few young enthusiasts, most scholars are still noticeably intimidated when faced with such tasks. This assessment holds not only for works on Nostratic

theory, but also for several other synthetic endeavors in comparative-historical linguistics (such as E. Matison's Amerindian hypothesis and Joseph Greenberg's Indo-Pacific hypothesis).

It seems that the validity of each of these concepts, as far as having some paramount value for an early ethnic history of mankind is concerned, cannot be based on lexical data alone—as was clearly understood by Illič-Svityč himself. If one operates with lexical correspondences alone—correspondences that include nouns and verbs, but not pronouns—then one runs into exceptional difficulties when one tries to separate the common, inherited lexicon of the languages being compared from a stock of common borrowings. As a clear illustration of this, one may cite the correspondence between IE *\*n(ō)m-n̥* 'name' and genetically related forms in Finno-Ugric and Samoyedic. While indicating that this set is restricted to two language families, Illič-Svityč states (p. 83) that any borrowing hypothesis seems to be excluded because of their phonological shape. Earlier, Pedersen (1950:103-4) had observed the striking similarity between Toch. *ñem* and the Finno-Ugric forms. This observation, as well as the probability of early Uralic-Tocharian connections, makes Illič-Svityč's conclusion questionable. Another typical example is the correspondence between IE *\*mer-* 'young man' and the Kartvelian etymon, questioned by Illič-Svityč himself (p. 40). However, as not only Hurrian (p. 39), but also Urartian, *mari* is a loan from Mitannian Aryan *marya-*, one is permitted to interpret dialectal Svan *mā re* 'person, man', Cecen *mār* 'husband' (p. 40) in the same way, namely, as an archaic loan from an Aryan source in the languages of the Caucasus, see Diakonoff (1963:39, 81, 85, 89) and Melikišvili (1971:231, No. 3; 286, No. 4).

Not only do particularly complicated lexical problems arise in cases in which a correspondence appears in but two of the six language families compared, but also when there is a broad zone of correspondences surrounding one of those two families. Cf., similarly, (reduplicated?) Sem. *šmš* 'sun' ('sun god' in Akkadian), Berber and Chadic *\*mš* 'fire, sun, morning, star' on the one hand and Kartv. *\*mž-e* 'sun' on the other hand. This is a unique instance of Kartv. *\*ž* < Nostr. *\*ž*, (p. 78). However, this word also appears in Abkhazo-Adyghan as *\*m(a)za-* 'moon',

so Klimov (1969:288), as well as in Daghestanian \*ba<sub>3</sub>a<sup>1</sup> 'moon, new moon' with correspondences in Nakh (Nakh-Daghest. \*b̄a<sub>3</sub>-), see Trubetzkoy (1928:165). In other words, this lexeme is common to all Caucasian languages (though semantically the Kartvelian forms display a unity with those from Semitic, and they thus stand in opposition to the North Caucasian and, in part, the Berber-Chadic forms). I do not want to make any final conclusions about the general character of this set of correspondences, but I do want to point out that, in the early periods of human culture, words designating moon and sun, as well as their corresponding deities, could very well present themselves as migratory terms. It is, nevertheless, worth noting that, for these very families, namely Semito-Hamitic and Kartvelian, a different word for 'sun' is assumed: Egyptian r<sub>9</sub> 'sun god' vs. dialectal Svan r̄ahi 'dawn' (p. 106). As for possible reflexes of this (borrowed?) stem in archaic Indo-European languages, cf. Luwian māšš-ana- 'god', etc.

As underscored by Illič-Svityč on several occasions, phonetic correspondences are crucial for the elimination of loans. One can say that the greater the number of phonetic transcodings and, accordingly, the greater the complexity of sound changes from Nostratic to the individual language families (which means that the connections between the forms being compared become less evident), the more secure the assumption about the existence of a mutual, common Nostratic source. This implies that the forms are inherited, not borrowed. In these respects, those words are particularly illustrative that, according to Illič-Svityč, reveal a simplification of sonorant clusters and vocalism, as well as a restructuring of roots in Indo-European, Semito-Hamitic, and Kartvelian—in other words, in the Western Nostratic languages (see p. 19 *et passim*). Another important group of phonetic transformations, broadly represented in the volume reviewed here, shows the evolution of initial palatal sonorants: \*ñ-, \*ñ̃-, \*λ-, etc. As unexpected as it may seem at first glance, correspondences between stems in j- in Indo-European and those with n- in other Nostratic languages as proposed by Illič-Svityč (p. 90 ff.) are more plausible than some superficial and deceptive correspondences. A Nostratic linguistics will make further progress only when and if

more rules of this kind are discovered, that is, rules which transcode forms of language A into forms of language B with a sufficient degree of complexity, though this will render credible Nostratic etymologies less attainable, but, simultaneously, more reliable.

Among those phonetic correspondences in this second volume which have particular relevance for the binary comparison of individual languages, the following is worth singling out for attention: IE *\*-H-* vs. Uralic *\*-k-* (p. 97). Additional probative examples have been found for this correspondence in recent years when genetically related grammatical (verbal and nominal) desinences were identified in Indo-European and Finno-Ugric, see Čop (1974:42-3).

The consistent identification of all grammatical and ancillary morphemes, including pronouns, in Nostratic languages seems to represent the principal means of justifying Nostratic theory. Here, I would like to emphasize that a review of Illič-Svityč's work compels one to conclude that the precise reconstruction of elements with syntactic functions is especially relevant; such elements may become strictly morphological only in the later history of individual language families.<sup>2</sup> Illič-Svityč clearly demonstrated the character of such elements on the basis of *\*mA* in relative constructions which "preserved the old status of the relative element in Semito-Hamitic where traces of its use as both a preposition and a postposition are revealed" (p. 47). For such particles, both a broad functional spectrum and possibilities of different positions in the sentence are typical; they later became prefixes in Kartvelian, whereas in other Nostratic languages the prepositional was generalized by transformation as a suffixal element (p. 48). In addition to the nominal and verbal elements described in this fashion by Illič-Svityč, one can mention similar functions for other verbal elements as well (e.g., Indo-European suffixal *\*-H-* vs. a corresponding prefix in Kartvelian).

In such cases it is possible, in at least some of the languages, to find traces of autonomous, independent usage of a particle which has not yet become an affix. Cf. Sem.-Ham. *\*la-*, *\*li-* (prepositional 'to') which can be compared with the particle *-la* in forms such as Svan *Ädiši-la* 'man from Ädiši' (p. 25; cf. Indo-European usage

as shown by Old Hittite *našili* 'in the language of the city of Nesa', *Hattušili* 'a man from Hattuš-', etc.).

A fragmentary outline of Nostratic nominal morphology is given in several articles in this volume. The formation of the heteroclitic paradigms in Indo-European and Dravidian has been explained by Illič-Svityč as an evolutionary result of combinations of oblique forms in *-n* with locative particles (of the type Nostr. *\*da*, *\*K'a*). Here one can point to the formation of complex nominal paradigms in East Nostratic as well as the radical restructuring of declensions in Semito-Hamitic (p. 81). Further study will enable us to specify this outline of the syntactic prehistory of the constructions from which nominal declensional forms evolved in the individual Nostratic languages. In order to comprehend the prehistory of declension in Nostratic languages, Illič-Svityč's observations about the development of *\*-mA* are particularly interesting. This formant for the marked direct object became the ending of the accusative singular (*\*-m*) in Indo-European (pp. 48-51). I note with satisfaction that Illič-Svityč (p. 79) shares my opinion of the interpretation of the Tocharian oblique case marker *-m* (< *\*-n*) as a trace of the Indo-European heteroclitic element *-n-*, as in Skt. *yak-n-ās* (gen. of *yak-r-t*) 'liver'.<sup>3</sup>

An especially fascinating correspondence with Kartv. *\*men-* (and genetically related forms in Uralic and Altaic, both East Nostratic languages) is the Indo-European stem of the principal oblique cases of the first person pronoun, analyzed by Illič-Svityč as *\*me-ne-* (pp. 63-6, 78-81). This analysis convincingly demonstrates the utmost importance of the corresponding pronominal paradigm (*\*mi*, obl. *\*me-ne-*, cf. 1st pers. pl. inclusive *\*mä*) for the validity of Nostratic theory in general. A comparable entry describing the 2nd pers. sg. pron. *\*t'i* can be pieced together on the basis of other entries (p. 80 ff.). One has to bear in mind that, in and of itself, the reliability of these reconstructed pronominal elements, as well as that of the other grammatical elements mentioned above, automatically presupposes the existence of a genetic relationship between Nostratic languages, a presupposition formed independently of the degree of acceptability of the etymologies of certain nominal or verbal stems. Possessed with the highly critical mind that he was, Illič-Svityč was himself ever willing

to check such etymologies and urged other scholars to check and recheck them in turn. He was always willing to doubt a substantial portion of his results, no matter how many concrete etymologies might be eliminated because of the non-native character of the correspondences from which they had been inferred or because of the purely sound-symbolic, that is, "descriptive" character of the stems being compared. Most instances of this sort were indicated by the author of the dictionary himself; but today, after the publication of two volumes of the dictionary, we can say that Illič-Svityč proved the existence of a genetic relationship between the Nostratic languages and proved it without a doubt. Both the necessary and sufficient conditions for such a conclusion have been dictated by reliable examination of the stablest elements of the grammatical system (including personal pronouns) as to their mutual relationship.

One of the most successful tests of Illič-Svityč's ideas appears to be the reconstruction of the system of inclusive and exclusive pronouns. Even Bopp studied the inclusive-exclusive category in Kartvelian in connection with some facets of Indo-European grammar, though his ideas found no support at the time, see Dondua (1975:75-95). When first Jensen (1930:117-26) and then Watkins<sup>4</sup> proposed that the Indo-European stem of the first person pronoun *\*n(e)-* had functioned as an exclusive, then these scholars themselves considered that their hypothesis was "at best unprovable," see Watkins (1969:47). However, the discovery (made simultaneously and independently by several scholars) of exclusive *na-* in Kartvelian<sup>5</sup> and of *n-* in the Chadic group of the Semito-Hamitic languages (so Pilščíkova 1959:217-8) permitted an alternative solution to the problem, a solution that corresponds to the tentative conclusion drawn earlier by the very Indo-Europeanists who were unable to accept it. According to Illič-Svityč: "The highly archaic inclusive function is preserved in Semito-Hamitic and can thence be reconstructed for Kartvelian. As for the Dravidian and Altaic data that present a different distribution of exclusive and inclusive forms, they can be explained as the result of a restructuring of corresponding systems in conjunction with new forms derived from the first person singular pronoun. In this way, the inclusive pronominal stem *\*mā* was opposed,

at the Nostratic level, to the exclusive pronoun of the first person *\*nV*." (p. 56). This example shows that Nostratic theory reflects a gradual development of Indo-European and comparative linguistics, as well as the comparative linguistics of other language families such as Kartvelian; the initiation of this process goes back as far as Bopp. The particularly important results of Nostratic theory lie in those points that, without the application of external comparison, would otherwise be stalemated.

Another example that reveals how natural the transition is from the study of relic pronouns in individual language families (such as Indo-European) to more encompassing Nostratic comparisons is as follows: Hittite *ma-* (in *maši-* 'how many') and corresponding forms in Tocharian and Celtic vs. the genetically related interrogative pronouns in Semito-Hamitic and Uralic to which a recent investigator added Kartvelian and Dravidian parallels (pp. 66-8). It was, after all, Pedersen who first suggested a genetic relationship between Semito-Hamitic, Uralic, and Indo-European forms (after the latter had been properly uncovered), see Pedersen (1938:67-72).

From among the various entries that are of considerable interest for the reconstruction of a Nostratic grammatical system, that treating the reconstruction of prohibitive *\*mā* (pp. 56-7) is particularly noteworthy. This particle is present in all families except Uralic (where the functions of the prohibitive negative are taken over by Uralic *\*ela*, *\*älä* < Nostr. *\*ʔälä*): the functional relationship between Nostr. *\*mā* and *\*ʔälä* is identical to that between IE *\*me* and *\*ne* (a possible trace of *\*ʔälä* can be seen not only in Hitt. *le*, but also in Slavic *\*li*, which has been reinterpreted in accordance with the well-known dictum that negative particles are transformed into interrogative particles). For Uralic, see Čop (1974), and for Indo-European see Lehmann (1974:153-4), cf. Ivanov (1977). Illič-Svityč's etymological hypotheses are highly significant, not only for a comparative grammar of Nostratic in general, but also for an understanding of the history of certain words in individual language families, such as the names of body parts in the various Indo-European dialects: 'spleen' (> IE *\*l[e]p-* 'liver' in Armenian and Greek, p. 17); 'leg' in Germanic (IE *\*lak-*, pp. 22-3); 'paw' in Slavic (IE *\*lep-*, pp. 23-4); 'scrotum' in Slavic (*\*mādo*, p. 73), etc. The

hypothesis that the value 'sea' for the IE radical *\*mor-* is secondary can be considered as finally settled because of the semantic uniformity of this Nostratic stem ('water', 'liquid', 'body of water') in all other Nostratic languages (except Middle Mongolian 'sea', pp. 60-1). This value is essential when considering Indo-European migrations. In this dictionary, specialists of Kartvelian languages will find a whole series of corroborations of the archaic character of Svan (cf., eg., p. 33, concerning the words for 'water' *lic-*, *nic-*). Of the various fine observations on semantic properties, one may single out the remarks about the similar revolutionary careers of IE *\*lip-* in Tocharian and Icelandic ('to adhere, stick' > 'remain, stay' > 'be superfluous, extra'), p. 18; of IE *\*lem-* in Albanian and Slavic ('to thresh grain' > 'threshing floor'), p. 22; of IE *\*lep-* in Celtic and Slavic ('shovel' > 'oar'), p. 23; and so on. Slavic semantic parallels are also successfully employed to explain the evolution of Toch. A *ljǝšk* 'groin' (< *\*les-* 'weak', cf. Bulg. *slabiní* 'groin', p. 16); of Drav. *nīvu* ('to smear' > 'to stroke'), p. 19, and so on.

Of those topics that are but cursorily mentioned in the dictionary, but which merit substantial attention, one might point to the possibility of retrieving traces of extinct Nostratic languages in loans, e.g., in Uralic (p. 45).

As soon as the number of languages that are compared is increased, and this number was deliberately restricted by Illič-Svityč, one may well be able to find additional corroboration for certain details of the many reconstructions proposed in the dictionary. Here, for example, it is sufficient to mention the function of the particle *-l* in Yukaghiran in determining the role of corresponding Nostratic forms (pp. 20-1), see further Kreinovič (1958:106 f.), Rosenkranz (1966:177).

In several instances the dictionary highlights the significance of archaic Hittite forms for Nostratic reconstructions (Hittite often preserves archaic phenomena lost in most Indo-European languages), cf. for example, Hittite forms of the type *tariyalla-*, *tarriyanalli-* 'third' vs. collectives in *-l* in other Nostratic languages (including collectives derived from numerals, as in Uzbek *ūčā-la* 'all three, the three of them, three together', p. 14). As far as the lexicon is concerned and items that should have

been included in the second volume of the dictionary, one could mention Hittite *neka-* 'sister, daughter, younger female relative' vs. Altaic forms of the type Tungusian (Neghidal) *nəxú(n)* 'younger brother, younger sister,' etc.

It is worth mentioning (as was done in reviewing the first volume) that a very precise, scholarly, and conscientious preparatory job of editorial work was done by V. A. Dybo, as well as by all the other scholars who helped him edit this second volume. The indices to both volumes, compiled by R. V. Bulatova and published at the end of the second volume, are invaluable.<sup>6</sup> Here, I would like to wish the whole team of scholars continuing the work of Illič-Svityč a successful completion of the third volume that is to be pulled together from the file cards that remained in the archives of the late founder of Nostratic linguistics.

## NOTES

### Ivanov/Review II

1. See Gigineišvili (1977:75, 84). For the initial *m-*, cf. Avar *moč*<sup>9</sup> 'new moon'; secondary *-b-* is represented in Kartvelian by Laz *bžačxa* 'Sunday', Megrel *bža* (p. 78).
2. On this tendency in the evolution of the Indo-European languages, now see Ivanov (1977:20-24).
3. For a more detailed presentation of this thesis, see Ivanov (1959:24-7).
4. See Watkins (1969:47), where, it appears, that author was unaware of Jensen (1930).
5. See Gamkrelidze (1959:46-7). Earlier proposals are much less clear, see Dondua (1975:82, 83, 92).
6. An important supplement is provided by R. V. Bulatova (1977:74-9).

B. A. Serebrennikov, *Problema dostatočnoisti osnovaniia v gipotezax, kasaiušćir'a genetičiskogo rodstva ĭazykov. 3. Nostratičeskie ĭazyki. In: Theoretical Foundations for the Classification of the World's Language. Problems of Genetic Relationship. Moscow, 1982. Pp. 47-62*

### Nostratic Languages<sup>1</sup>

The development of comparative-historical linguistics has always been accompanied by attempts to enlarge the circle of genetically related languages, mainly by the discovery of correspondences between two or more families in order to establish a genetic relationship for them inherited from the past.

The history of linguistics records several attempts to compare Indo-European with Semitic; Finno-Ugric with Altaic; Indo-European-Turkic with Uralo-Altaic, and so forth. Attempts have also been made to link Turkic languages genetically with Altaic; Turkic with Mongolian; Finno-Ugric with Dravidian, Uralic with Yukaghiran, and so on.

Previously, scholars tried to compare pairs of languages, such as Semitic and Indo-European, Turkic and Finno-Ugric, etc., but currently scholars are attempting to compare entire phyla, even those that consist of an extraordinary number of languages.

Morris Swadesh tried to forge a genetic connection between the languages of the Old World and those of the New. He suggested the existence of a large macrofamily, which he termed *Dene-Finnic (Finnodenean)*, after the names of the two geographically most widely separated members of this macrofamily: The Dene-Athapascan languages of North America and the Finno-Ugric languages of Eurasia.

Illich-Svityč identified a large macrofamily of languages that he labelled Nostratic. This macrofamily includes six large families of the Old World: Indo-European, Altaic, Uralic, Dravidian, Kartvelian, and Semito-Hamitic.

The desire to identify additional related languages seems to be a natural one. The history of linguistics records instances of such identification that were crowned with success. At the outset, the Samoyedic languages were

considered unrelated to Finno-Ugric, but after the Finnish linguist, E. Setälä, published his famous work, *Zur Frage nach der Verwandtschaft der finnischugrischen und samoyedischen Sprachen*, and conducted subsequent investigations, the genetic relationship between Samoyedic and Finno-Ugric was finally established; and now Samoyedic data are widely used by scholars of Finno-Ugric. This demonstrates the fact that the precise genetic relationships between and among the various languages of the world have not yet been fully established and that the possibility exists that yet other languages can be unified genetically with a given group of languages whose genetic affiliation has been sufficiently established.

As noted above, current linguistic investigation is attempting to establish macro-phyla that include an extraordinarily large number of languages. This is not merely a flight into the fantastic, but a purposeful endeavor with certain principles:

(1) The unification of a vast number of languages from different families into one macro-family with one ancestor for these languages broadens the framework of historical and developmental perspectives enormously. Thus, if the Uralic proto-language existed some 10-15 thousand years ago, then the Nostratic protolanguage would have to have existed much earlier. It is possible that it was extant some 30 or even 40 thousand years ago.<sup>2</sup>

(2) The greater the number of genetically related languages in a given family, the more probable the preservation of some exceedingly ancient archaisms. The study of these could help in many ways to render our knowledge of the history of modern languages more precise and exacting. The discovery of new phonological laws and some archaic elements of grammatical structure could contribute to a more accurate understanding of the general course of evolution in genetically related languages.

(3) The discovery of large macrofamilies of languages could contribute to a more refined definition of the geographical displacement of related languages in the distant past.

A rather intriguing linguistic problem arises from all this: Is genetic relationship ever really proven, and, if so, how can one verify it? This is an extremely difficult question.

We cannot say that, when establishing the genetic relationship of Nostratic languages, Illič-Svityč ignored phonetic correspondences, for at every step he certainly cited correspondences, whose establishment was periodically accompanied by genuine virtuosity. It is much more difficult to examine and verify these correspondences than others because they extend over a vast number of languages that are exceedingly diverse in structure. Then, too, one might add that the histories of some of the families, such as Hamitic or Kartvelian, have not yet been thoroughly investigated. Therefore, we have selected another means by which to evaluate the Nostratic hypothesis. It seems advisable to test the general effectiveness of this hypothesis by first singling out some of the Nostratic languages.

Along these lines, one may formulate the following question: "What bearing does the Nostratic hypothesis have on the history of Finnish?" The fact is that Finnish was included in Finno-Ugric; and historical-comparative investigation permitted, in turn, linguists to reconstruct the history of Finnish. After it had been proven that Samoyedic languages were, in fact, genetically related to Finno-Ugric languages, Finnish was considered a Uralic language.

The fact that Samoyedic was added to Finno-Ugric did not, however, trigger a revolution within Finno-Ugric comparative studies. The vocalic and consonantal structures of the Uralic proto-language as previously retrieved remained virtually intact, and no especially archaic grammatical forms have been discovered. Samoyedic languages contain a large number of innovations, and Finno-Ugrists frequently encounter considerable difficulties when comparing Samoyedic data with those of Finno-Ugric.

According to Nostratic theory, Finnish belongs to Nostratic, and this implies that it is not only related to Finno-Ugric and Samoyedic, but also to Turkic, Mongolian, Dravidian, and Semito-Hamitic. In this context, it would be interesting to discover to what extent Nostratic theory might enrich and deepen our knowledge of the history of Finnish.

One of Illič-Svityč's incontrovertible contributions consists in the fact that he assembled all the phonetically similar lexemes of Finno-Ugric, Indo-European, Turkic, Mongolian, Dravidian, Kartvelian, and Semito-Hamitic.

This was accomplished on the basis of work by his predecessors as well as from his own observations. As a major result of this detailed and intricate compilation, a comparative dictionary of the Nostratic languages appeared in two volumes which Illič-Svityč (1971, 1976) entitled *Studies in the Comparison of the Nostratic Languages*.<sup>3</sup>

It is well known that, insofar as the genetic relationship of languages may be determined, the material relationship of grammatical formants is more relevant than a comparison of lexical roots.<sup>4</sup> In the lexicon, occasional coincidences and loans can occur. It is also well known that inflectional elements are never borrowed; but as for the constituents of derivation, of word formation, they are borrowed only if languages are in intensive contact.

In his dictionary, Illič-Svityč not only compared words, but also formants; but data for the Nostratic case system are rather limited. In Nostratic, there was a marker *-n* for the oblique form of nouns and pronouns. In Indo-European, this state of affairs was preserved in nominal heteroclis, that is, in forms such as Skt. *yak-r̥-t* 'liver', gen. sg. *yak-na-s*; Lat. *femur* 'thigh', gen. sg. *feminis*; Russ. *ja* 'I', gen. *men'a*, Finn. *minu-a*, 'of me, mine' (from the stem *-mi*), cf. Chuvash *ebě* 'I', gen. *man-ăn*, etc.

After the evolution of a more complex declension system, the functional load of forms in *-n* was reduced: it has been preserved only as a variant of the genitive.<sup>5</sup>

Indeed, a certain similarity exists between the endings of the genitive in Nostratic languages, cf. Finn. *kala* 'fish', gen. sg. *kala-n*; Literary Mongolian *gar* 'hand', gen. sg. *gar-un*; Old Tamil *ūr* 'village', gen. sg. *ūr-in*. In Turkic languages, as a result of the merger of a primary ending in *-n* with an adjective suffix *-ki*, a genitive marker *-yn*, *-in* evolved.

Such an explanation of the origin of the genitive in Finno-Ugric and Turkic is rather contradictory.<sup>6</sup> Oblique pronominal stems containing *n* already existed in Turkic at the time when the genitive ending had not yet come into being. An *izafet*-construction was used for the genitive, as in Yakut *at baha* 'horse head'. A genitive suffix emerged in Turkic as a result of the use of the instructive in this sense. After the instructive ending was employed as a genitive, the use of the instructive was greatly reduced

in Turkic. In modern Turkic languages, only a handful of adverbs preserve remnants of the instructive: Turkish *yaz-in* 'in summer', *kiş-in* 'in winter', and so on.

In Finno-Ugric, too, it appears that the genitive ending was originally that of an instructive, cf. Mari *jal-ən* 'of the village', *jol-ən* 'by foot'. After the elimination of coaffixal *-l-* in the Komi genitive-*lõn*, *-õn* remained, which coincided with the ending of the instrumental, as in *êcer-õn* 'with (an) axe'.

In the Nostratic proto-language, according to Illič-Svityč, there was a formant for the marked direct object *\*-mA*. In the successor Nostratic languages, this formative is preserved as the accusative, e.g., in Skt. *vrka-m* 'wolf', Lat. *lupu-m* id., cf. Finn. *kala-n* 'fish' (< *\*kala-m*); as for Tungusian, the accusative is represented by *-ba/-bä*, as well as *ma/mä* (after nasals). Traces of this ending are also found in Dravidian. However, this explanation is also contradictory.<sup>7</sup> In Indo-European, the accusative seems to have originated from some sort of lative case, cf. Skt. *graman gachati* 'goes to the village', Lat. *Rōman īre* 'go to Rome'. This grammatical significance is absent in Finno-Ugric. One cannot correlate the Tungusian accusative ending *\*-ba/\*bä* with an Indo-European or Finno-Ugric ending of this case in *\*-m/*

According to Illič-Svityč, Nostratic had a locative particle *\*-da*, which was originally used as an enclitic, mostly with pronouns and other stems that indicated spatial deixis. It was only later that this particle was used as a postpositional after nouns and finally became a case formant, cf. Illič-Svityč (1971:212-4). Again according to Illič-Svityč, the Indo-European and Uralic ablative formants originated from this locative particle, cf. Skt. *ma-t* 'from me', Osc. *tōut-ad* 'from the community, people', Finn. partitive *talo-a* < *\*talo-δa* 'from the house', as well as the ablative formant *\*-ttu*, *\*-du* in Dravidian and the ablative-locative formant *-tā* in old(er) stages of the Turkic languages.

However, marked differences in the semantics of the ablative in various Nostratic languages cast doubt on the genetic identity of this formant in Nostratic. In Indo-European and Uralic, this formative bears ablative significance. In the Turkic proto-language it had both locative and ablative meaning, cf. Orkhon *taγ-da* 'on the mountain',

but *ga ɣan-ta* 'from the emperor'; in Mongolian this formant had both locative and lative significance, cf. Buryat *bulan-da*, both 'in the corner' and 'into the corner'. In Semito-Hamitic and Kartvelian, this formant has lative significance, but neither ablative nor lative meaning is derivable from a locative.<sup>8</sup>

It is a well-established fact of Balto-Finnic that coaffixal *l* is present in the series of extra-local cases, cf. Finn. *rannal-la* 'on the shore', from *ranta-l-na*; *talolla* 'on the house', and so on.

This coaffixal element also appears in Mari and Permian, cf. Mari *čodra-l-an*, Komi-Zyryan *vör-l-añ* 'toward the forest'.

In line with Illič-Svityč's view, Proto-Uralic had a locative suffix used in the formation of *nomina loci* of the Finn. *etelä* 'south' type. Illič-Svityč connects this suffix to that of the locative *-la/-lä* in Dravidian, cf. Evenki and Even *mō-lā* 'on the tree'. As for Semito-Hamitic, this element had already been established as a preposition, *la*, *li*, or *l*, with the meaning 'toward'. Nostratic is presumed to have had a particle *\*K'V* with lative significance, cf. Illič-Svityč (1971:368), which is usually reflected by the lative ending, as in dial. Finn. *ala-k* 'down', Mari *vel-ke* 'to the side', Tamil *ūru-kku* 'to, into the village', Tatar *jaq-qa* 'to the side', Evenki *anti-kā* 'to the right', and so on. There is reason to believe that the lative ending, *-qa*, in Turkic is comprised of two lative suffixes, namely *-q* and *-a*, though the latter lacks any correspondence in Nostratic.

Nostratic had a notion of plurality, though it was treated as a collective plural, and several collective particles can be identified:

The particle *\*q'V* is reflected by the Armenian plural *-k'*, e.g., Arm. *bank'* 'words'. In Uralic, this *\*q'* is reflected in the personal endings of verbal plurals, cf. Finn. *annamme* 'we give' < *annammek*. A reflex of this same particle is also found in Dravidian: Gondhi *kai-k* 'hands', *tele-k* 'heads', cf. Illič-Svityč (1976:105).

The marker *\*-l(a)-* is found in Turkic as part of the plural suffix *-lar*, cf. Turkish *at-lar* 'horses'. Mari attests a collective suffix *-la*, as in *kož-la* 'small fir-tree forest'. In Dravidian, one also finds suffixal *\*-l*, e.g., Komami *kaje-l* 'hands', cf. Illič-Svityč (1976:14).

For the suffix *\*-na*, cf. Arabic *-an* as in *furs-an* 'horsemen' from *faris* 'horseman'; old Turkic *oɣl-an* 'children'

from *oγul* 'child'. In Uralic, traces of this suffix are present in declensions with the possessive suffix, cf. Erzya-Mordvinian *kede-n-ze* 'his hands', cf. Illič-Svityč (1976:94-5).

For Nostratic, a nominal suffix *\*j(V)* is posited for the (plural) oblique form, cf. Illič-Svityč (1971:285). In Indo-European, it is reflected in the oblique forms of demonstrative pronouns, cf. Gk. *τοί* 'these', Goth. *þai* as in *o*-stem nouns, e.g. Gk. *λύκοι* 'wolves'. This suffix can also be found in Finnish, e.g., *talo-i-ssa* 'in the houses'.

One should emphasize the fact that the reconstruction of a Nostratic proto-language by no means reveals the total richness of suffixes denoting collective plurality that must have existed at some time or other in Uralic.<sup>9</sup> There were at least ten suffixes in this category: *-a* (*-ja*), *-ĉ*, *-i* (*-j*), *-k*, *-l*, *-m*, *-n*, *-r*, *-s*, *-t*. Each of them, at least initially, designated a collective plurality for a certain noun class that differed from similar classes by specific markers. However, the Indo-European and Semito-Hamitic suffixes, while phonologically similar, do not reveal this feature of noun-class inclusion.<sup>10</sup>

There are several phonologically similar nominal and verbal suffixes:

Suffixal *\*-jV* with diminutive-endearament significance, cf. Gk. *παῖδ-ιον* 'little child' < *παῖδ-* 'child', Tatar *ata-j* 'little father' < *ata* 'father', Mengrel *baba-ja* 'little father' < *baba* 'father', Finn. dial. *emo-i* 'little mother', see Illič-Svityč (1971:284).

The suffix *\*-k'a* with diminutive significance (cf. Illič-Svityč 1971:312), as in Georgian *mam-ik' -o* 'little father', Finn. *vasi-kka* 'little calf' *vasa* 'fawn', Old Uigur *ata-qy* 'little father', Skt. *aśva-ka* 'little horse', Russ. *mýš-ka* 'little mouse', and so on.

The suffix of deverbative nouns *\*-mA*, (cf. Illič-Svityč 1976:45), as in Finn. *juo-ma* 'drink', Turk. *jaz-ma* 'writing', and so on.

The suffix of denominative verbs *\*-l* (cf. Illič-Svityč 1976:16), as in Erzya-Mordv. *ekše-la-* 'to bathe' < *ekše-* 'cool', Tatar *baš-la* 'to begin'

*baš-* 'head', even *mū-la-* 'to carry water' *mū* 'water'.

The suffix of frequentative verbs *\*-Ĉi* as in Finn. *ruoski-tse-* 'to whip', Erzya-Mordv. *mor-šə-* 'to sing (often)', Nanay *xola-č-i-* 'to continue reading', Hitt. *ar-šk-* 'to move (something) often', and so on, cf. Illič-Svityč (1971:206).

Here again, the proposed reconstruction of Nostratic does not fully reveal the entire wealth of frequentative suffixes in the Uralic languages.<sup>11</sup> There were about ten such suffixes in Uralic that are genetically related to those of the collective plural.

There are only highly restricted data concerning the tense-aspect system of Nostratic. Only one past tense is reconstructed: namely, that with the marker *\*-di*. The marker *\*-j(-i)*, typical for Uralic, was presumably present in Nostratic as a particle that could be placed either before (e.g., Gk.  $\epsilon-\gamma\epsilon\rho\epsilon$  'he carried') or after (e.g., Finn. *tul-i* 'he came') a verbal stem, cf. Illič-Svityč (1971:249).

This correspondence is not very persuasive.<sup>12</sup> In Greek, the augment seems to have originated from some adverbial marker. As for the preterital marker *\*-j* in Uralic, its origin is not yet entirely clear. Moreover, there are no parallels for this tense-marking strategy in Turkic, Mongolian, or Dravidian.<sup>13</sup> Then, too, there was a past tense in *\*-š* in Uralic that lacks Nostratic parallels, though this marker apparently antedated *\*-j*.

According to Illič-Svityč, Nostratic had an optative *\*je* which became *\*-i-* in Indo-European (thematic stems displayed *\*-iē-* as in Gk.  $\gamma\epsilon\rho\omicron\iota$  'let (him) carry'.) Illič-Svityč connects this optative with that in Turkic (*\*-aj/\*-aj*) and Mongolian (*\*-ja/\*-je*), see Illič-Svityč (1971:282).

The category of mood makes its appearance late in the various languages of the world. The Indo-European injunctive represents the indicative of the aorist. The marker of the conditional-desiderative in Finnish (*-ne*) originated as a result of reinterpretation of the semantics of a frequentative suffix. The Turkic optative marker *\*-aj* (Turkish *yaz-ay-im* 'let me write', etc.) is genetically identical with the diminutive suffix *\*-aj* (as in Tatar *an-aj* 'little mother'). Modality in Turkish originated as a result

of reinterpretation of the semantics of diminutive suffixes. It strains credibility to assume that the Nostratic proto-language, which originated some 40 thousand years ago, possessed an optative.<sup>14</sup>

The comparative Nostratic dictionary also lists some similar forms for various pronouns, but these had already received scholarly attention. From the above, one can see that the data for Nostratic are rather attenuated; and sporadic correspondences are frequently not very persuasive;<sup>15</sup> and they therefore provide no information that is crucial for the history of Finnish.<sup>16</sup>

All the grammatical parallels cited by Illič-Svityč had, in the main, been known previously, but no one was able to employ them to uncover the specifics of Uralic grammatical structure in the distant past.<sup>17</sup>

Let us now consider the following problem: what, for example, can a reconstruction of Nostratic phonetics offer the history of Finnish?

For the most part, Nostratic vowels remain unchanged:

Nostr. *\*a* > Finn. *a*, cf. Nostr. *\*Haja* 'chase' > Finn. *aja-*; Nostr. *\*kalV* 'fish' > Finn. *kala*, Nostr. *\*marja* 'berry' > Finn. *marja*.

Nostr. *\*e* > Finn. *e*, cf. Nostr. *\*kerjä* 'shout' > Finn. *kerjää-* 'beg'; Nostr. *\*?ela* 'live' > Finn. *elä-*; Nostr. *\*pelHi* 'to fear' > Finn. *pelkä-ä*.

Nostr. *\*ä* > Finn. *ä*, cf. Nostr. *\*gāt'i* 'hand' > Finn. *käte-*; Nostr. *\*k'ärV-* 'to bind tightly' > Finn. *kääri-*.

Nostr. *\*o* > Finn. *o*, cf. Nostr. *\*Koki-* 'to trail, to observe' > Finn. *koke-* 'to try'; Nostr. *\*HonČa* 'end, edge' > Finn. *otsa* 'forehead'.

Nostr. *\*u* > Finn. *u*, cf. Nostr. *\*duli* 'fire' > Finn. *tule-*; Nostr. *\*burV* 'snowstorm, sandstorm' > Finn. *purkuu* 'snowstorm'.

Nostr. *\*ü* > Finn. *ü*, cf. Nostr. *\*Küt'V* 'bind' > Finn. *kytke-* 'harness'; Nostr. *\*kal'iv* '(to get) cold' > Finn. *kylmä* 'cold'.

Nostr. *\*i* > Finn. *i*, cf. Nostr. *\*kiwi* 'stone' > Finn. *kivi*; Nostr. *\*kirV* 'hoarfrost' > Finn. *kirte-*.

There are some instances of incomplete coincidence: as, for example, Nostr. *\*a* vs. Finn. *o*; Nostr. *i* vs. Finn. *a*; Nostr. *\*e* vs. Finn. *a*; but no explanations for these discrepancies are offered. In general, however, these sporadic deviations do not alter the overall picture—that of a good presentation of Nostratic vocalism in Finnish.<sup>18</sup> Here, we have a spectacular miracle in linguistic history. Nostratic was extant some forty thousand years ago, but across this immense expanse of time the vocalism of this proto-language was retained in Finnish without substantial changes.

According to Illič-Svityč's reconstruction, a typical feature of Nostratic consonantism was the presence of so-called "abruptives" and laryngeals, as well as affricates. None of these has been preserved in Finnish, cf. Nostr. *\*k'ajwa* 'dig' (with an abruptive *\*k'*) > Finn. *kaj* *a-*; Nostr. *\*?ela* 'live' (with laryngeal *\*?*) > Finn. *elä-*; Nostr. *\*Haja* 'to chase' > Finn. *aja-*. One assumes that Nostratic voiced stops underwent devoicing in Proto-Uralic, cf. Nostr. *\*duli* 'fire' > Finn. *tuli*; Nostr. *\*bara* 'big, good' > Finn. *paras* 'better'. A specific *l*-phoneme is posited for both Nostratic and Proto-Uralic that is reflected by *t* in Finnish, cf. Nostr. *\*lamHu* 'wild cherry' > Finn. *tuomi*; Nostr. *\*majλ V* 'sweet sap' > Finn. *maitē* 'milk'. Nostratic affricates have not been preserved in Finnish.

I doubt whether Nostratic had laryngeals.<sup>19</sup> Illič-Svityč posits a laryngeal whenever he requires an explanation for vowel lengthening, e.g., in Old Turkish *ūč* 'edge' vs. Finn. *otsa*. To explain old Turkish *ū*, Illič-Svityč postulates a Nostratic laryngeal: Nostr. *\*Honča* 'end'.

Now, if we pose the questions: does Nostratic theory provide anything new for the history of Finnish, does it enrich and deepen our knowledge of the history of Finnish?—then the answer would be negative.<sup>20</sup>

As for the details of Nostratic morphology, these, too, were recognized by previous scholars; but, as their interpretation was so vexed, Finno-Ugrists made practically no use of them. As for the consonantism, it is based on highly dubious hypotheses; and the vocalic correspondences

are a simple paradox: Nostratic vowels failed to undergo any substantial changes.<sup>21</sup>

Now, let us assume that our general conclusion is wrong: low returns from the Nostratic hypothesis for the history of Finnish is merely an exception to the general rule. However, Illič-Svityč could not even substantially enlighten the problematic genetic relationship between Turkic and Mongolian. The phonological correspondences between Turkic and Mongolian are extremely monotonous and not contrastively rewarding.<sup>22</sup> Turkic *\*a* corresponds to Mongolian *\*a* in many instances, cf. Pre-Turkic *\*k'aryn* 'belly': Literary Mongolian *garbith* 'belly, fat'; Turkic *\*kara* 'black': Mongolian *\*kara*; Kirghiz *aṇ* 'hole': Mongolian *aṇ* 'hole'; Old Uigur *galy-* 'to lift': Literary Mongolian *qali-* 'to glide'. Turkic *\*o* often corresponds to Mongol *\*o*, cf. Turkic *\*ok* 'arrow': Literary Mongol. *oki* 'top'; Tuvian *ɣol* 'dry river bed': Mongol. *\*gol* 'river'. Turkic *\*u* often corresponds to Mongol. *\*u*, as in Turkic *\*-i* (possessive affix of the 3rd pers. sg.): Mongol. *\*i* (demonstrative pronominal stem). Turkic *\*e* corresponds to Mongol. *\*e*, as in Old Uigur *keṣi* 'end, edge': Literary Mongol. *gede* 'back of the head'. Turkic *\*u* can correspond to Mongol. *\*ü*, cf. Turkic *bür-a-* 'to cover': Mongol. *\*büri-* id. Old Turkic *ä* generally corresponds to Mongol. *\*ä*, cf. Old Turkic *äm(a)-* 'to suckle': Literary Mongol. *em-kü-* id.; Old Turkic *är* 'man' ('male'): Mongol. *ere* id. Turkic *\*y* usually corresponds to Mongol. *\*i*, as in Old Turkic *k'yrālu* 'hoarfrost': Literary Mongol. *kiragu*; Turkic *k'yl* 'hair': Literary Mongol. *kili* 'string'. There are cases in which Turkic *\*u* corresponds to Mongol. *\*o*, cf. Turkic *\*kurt* 'worm': Literary Mongol. *gorogaj*. Turkic *\*e* can also correspond to Mongol. *\*i*, as in Turkic *\*elik* 'deer': Mongol. *\*il* id., though the causes of these deviations are unclear.

When comparing Turkic and Mongolian vowels with those of Nostratic, we again encounter a stunning paradox: many Nostratic vowels do not undergo any changes in either Turkic or Mongolian.<sup>23</sup> Thus, Nostr. *\*a* becomes Turkic and Mongol. *\*a*, e.g., Nostr. *\*Hanga* 'to open the mouth' > Kirghiz *aṇ* 'hole', Mongol. *aṇ* id.; Nostr. *\*k'ar(a)* 'black' Turkic *\*kara*, Mongol. *kara*, *χara*). Nostr. *\*o* corresponds to Turkic and Mongol. *\*o*, cf. Nostr. *\*Hok'i* 'sharp point' > Turkic *\*ok* 'arrow', Literary Mongol. *oki*

'top'; Nostr. *\*golHV* 'heart' > Tuv. *χol* 'dry river bed': Mongol. *\*gol* 'river'. Nostratic *\*u* corresponds to Turkic and Mongol. *\*u*, cf. Nostr. *\*k'udi* 'tail' > Turkic *\*kuδ uruk* 'tail': Literary Mongol. *qudurqa* 'strap'; Nostr. *\*Huk'a* 'eye, see' > Turkic *uk(a)-* 'to lift', Mongol. *\*uqa-* id., etc. Of course, these correspondences periodically run into incongruities, but the general course of correlations remains the same.

Here it should be noted that, when establishing correspondences between Turkic and Mongolian languages, Illič-Svityč uncritically repeats some of the same errors as were committed by Altaists.<sup>24</sup> Following R. Rumstedt and N. Poppe, he accepts the existence of two *r*-phonemes: an unmarked *\*r* and a palatalized *\*r̥* that is not dependent on vowel harmony. Hence, the correspondences Nostr. *\*HirV* 'to drag' > Tatar *yzan* 'furrow'; Nostr. *\*šehrâ* 'to be awake' > Uighur *sez-* 'to feel', Mongol. *sere-* 'to doubt'. As is the case with Altaists, he postulates the existence of a Proto-Altaic *\*l* that is independent of vowel harmony, cf. Nostr. *\*kalV* 'to bark' > Chagatay *gašqa* 'bald', Mongol. *\*χalžan* 'bald'; Nostr. *\*qalV* 'to top a mountain' > Old Uigur *aš-* 'to cross'. Again, as is the case with other Altaists, Illič-Svityč accepts the possibility of a shift of initial *d-* to *j-*, as in Nostr. *\*duli-* 'fire' Mongol. *\*dula-gan* 'warm', but Turkic *\*jyly*, *\*jylyγ* 'warm'; Nostr. *\*daka* 'to get closer' > Literary Mongol. *daga-* 'to follow', Old Turkic *jaγ* - 'to stick to, to pass'. As N. Poppe, so Illič-Svityč, too, believes in the possibility of a shift from *\*p* to *\*γ*, as in Nostr. *\*K'ap'a* 'to cover' Turkic *\*k'ap(a)*, but Mongol. *\*kaγ.a-* < *\*Kɪpa* 'to cover' Turkic *\*k'ap(a)*, but Mongol. *\*kaγ.a-* < *\*kapa* 'to cover, close'.

Illič-Svityč frequently employs onomatopoeic terms in his comparisons, as well as words rooted in other forms of sound symbolism. This is, of course, methodologically incorrect.<sup>25</sup> Here is an example of such a set: Nostr. *\*čap'(a)* 'to beat' > Turkic *\*capa-*, Literary Mongol. *čabči-* 'to cut'.

Thus, as far as both Turkic and Mongolian languages are concerned, unfortunate results emerge from an application of the Nostratic hypothesis.<sup>26</sup>

Indeed, all such results are, it seems, not merely random or otherwise exceptional, for the whole system of Illič-Svityč's arguments in favor of a genetic relationship

among "Nostratic" languages has substantial drawbacks.

*Incorrect Etymologies.* From the roster of incorrect etymologies, those based on a comparison of onomatopoeic words have to be cited. There are many such etymologies in *The Dictionary of the Nostratic Languages*. Here are a few examples: Finn. *pura* 'drill, chisel', Lat. *forare* 'drill', OHG *borōn* id. Akkad. *buru* '(water) spring' < Nostr. \**bura* 'to bore'; Finn. *sorottaa* 'to drip', Tamil *čor-* id., Mongol. Baoan *čur-* 'flow', Hung. *csorog* 'flow' < Nostr. \**čurV-* 'to drip'; Finn. *kurke, kurkki* 'crane', Arabic *kurlij*, OHG *kranuh*, Malayalam *kurijan* 'crane' < Nostr. \**karV/\*kurV* 'crane'; Finn. *ime-* 'to suck', Old Turkish *am-*, Literary Mongol. *em-kū-* < Nostr. \**H[E]mi* 'to suck, swallow'.<sup>27</sup>

It is well known that onomatopoeic forms cannot be employed to prove relationships between languages, for convergences between such words are also found in unrelated languages, cf. Finn. *sopottaa* : Russ. *šeptat'* 'whisper' : Rum. *șopti* id.; Gk. βατραχος 'frog' : Erzya-Mordv. *vatrakš*; Gk. κρόαξ 'raven' : Mari *korak* 'crow'; Chuvash *šor-* 'gurgle' : Komi Zyryan *šor* 'creek'; Finn. *puuskua* 'to blow' : Gk. πιδάω id., and so on.

*Highly Hypothetical or Improbable Comparisons.* Some of the correspondences proposed by Illič-Svityč are highly hypothetical. Gk. γίω 'to grow' has been compared with Skt. *bhū-* 'to be' (as in *abhūt* 'he became'), Lith. *būti* 'to be', as well as with Mongol. *bu* 'be' and Finn. *puu* 'tree'. Finno-Ugric languages display no indication at all that the root \**pū* meant 'to grow'. Finn. \**elä-* 'to live' has been compared with Tamil *il* 'house' and Old Turkish *dl* 'people'. A relationship between these etyma is dubious. Doubtful, too, is the comparison between Finn. *kylmä* 'cold' : Goth. *kald-s* id. : Tamil *kulir* 'cold (noun)' : Mongol. *köl-de-* 'get cold'. Finn. *käsi* 'hand' (stem *käte-*) is compared with Gk. χείρ 'hand' (< IE \**hes-r*). The shift of *t* > *s* before *i* is a purely Finnish (more precisely, Balto-Finnish) phenomenon and can scarcely be compared with a heteroclitlic declension in Indo-European. Doubtful, too, is the comparison between Finn. *kara* 'thorn, wooden nail, fish bone' : Buryat *gar-* 'go out'.<sup>28</sup>

But then, there are weak points in Nostratic theory itself.

First, the Nostratic character of Semito-Hamitic (Afro-Asiatic) languages is very doubtful.<sup>29</sup> It is simply

impossible to imagine how Nostratic roots with vowels and consonants could have been transformed into a Semitic root that is usually represented by two or three consonants,<sup>30</sup> cf. Nostr. *\*buŕa-* 'to boil' > Finn. *por-ise-* 'to bubble', Sem.-Ham. *\*ɔr* 'to boil'; Finn. *taka* 'back' (adj.): Sem.-Ham. *dk* 'close to'; Finn. *kivi* 'stone' : Chadic *kw*; Finn. *lipa* 'slippery' : Sem.-Ham. *\*p*, and so on. The phonetic bulk of Semitic roots is highly restricted: three, frequently only two, consonants. One and the same sound sequence can represent a multiplicity of meanings, as in *\*qr* 'to call' and 'ice', *mn* 'to think' and 'man'. Thus, we conclude that comparisons between Finnish and Semito-Hamitic etyma do not provide us with anything that is ultimately very useful.

A genetic relationship between Finnish and Kartvelian, particularly Georgian, is also doubtful.<sup>31</sup> Note the comparative evidence for Finnish and Georgian etyma: Finn. *teke-* 'to make' : Georgian *d(w)*; Finn. *ela-* 'to live': Old Georgian *er* 'people, army'; Finn. *kyty* 'brother-in-law': Georgian *kwis*; Finn. *sorotaa* 'to drip' : Old Georgian *cwar-*; Finn. *kivi* 'stone' : Georgian *kwa*, and so on.<sup>32</sup>

In several of the roots of etyma used to connect different Nostratic languages, sound sequences appear that have the following structural sequence: *C + V + Resonant (= R)* (that is: *r, l, m, n*). This type of root structure (*C-V-R*) is particularly frequent in the various languages of the world. In such instances, then, we risk discovering purely coincidental correspondences.

In the course of reconstructing Nostratic vocalism, Illič-Svityč clearly follows the Uralic pattern, but when reconstructing the consonantism, he follows Kartvelian and Semito-Hamitic patterns.<sup>33</sup> Accordingly, the numerous problems of Indo-European vocalism (apophony) and consonantism are deprived of satisfactory explanations in Illič-Svityč's dictionary.<sup>34</sup>

The dictionary contains many unresolved problems. One could, for example, question why it is that *q* in Turkic languages nearly always corresponds to *q* in Semitic languages;<sup>35</sup> why there was originally no vowel harmony in Uralic and Turkic languages;<sup>36</sup> what was the original typology of the Nostratic proto-language itself, and so on. Finally, then, the genetic relationship of the so-called "Nostratic" languages is insufficiently proven.<sup>37</sup>

## NOTES

**Serebrennikov**

1(E). We have omitted four previous sections of this article, including one on Caucasian languages and one on Altaic languages, both of which would now be considered quite obsolete in view of advances in comparative linguistics during the past decade or so, see Ivanov's article on proto-languages in this collection (pp. 1), as well as Ivanov (1983), and note the Foreword to this collection. Here, too, it should be pointed out that S.'s article contains many factual errors and numerous misprints; we have commented on the former and attempted to correct the latter.

2(E). This is a misrepresentation. Both Illič-Svityč and Dolgopolsky believed that Nostratic existed as a proto-language some 10,000 years ago. Accordingly, as a daughter language of Nostratic, the proto-language of Uralic would have to be placed into a much shorter period of time.

3(E). So far, only two volumes of Illič-Svityč's Nostratic dictionary have been published. The third volume, with seven additional entries, is expected to appear shortly, while the fourth volume, which is based entirely on Illič-Svityč's notes, will include a grammatical sketch of Nostratic.

4(E). This is not entirely correct. Note, for example, that the list of the fifteen most stable sememes in Dolgopolsky's article in this collection (pp. 33) contains both grammatical and lexical items.

5(E). This is a gross misrepresentation of Illič-Svityč's point of view, for he (1976:18-81, Entry #314) stated that Nostratic \*-n was a suffix of the oblique form of nouns and pronouns; it continued to have this function in the earlier stages of Indo-European, Uralic, and Dravidian, only later becoming a regular genitive ending, albeit independently in Uralic and Dravidian. In Kartvelian and the Altaic languages, -n preserved its original function as the suffix for an undifferentiated oblique form for both nouns and pronouns.

6(E). See fn. 5 above.

7(E). Again, this is a misrepresentation. According to Illič-Svityč (1976:48-51, Entry #285), the Nostratic formative *\*-mA* supplied the acc. sg. of animate nouns in Indo-European; the suffix of the "marked" objective sg. (*\*-m*) in Uralic; and the suffix of the marked objective (*\*-m*) in Dravidian; and the suffix of the marked/determined object *\*-ba/\*-bã* in Tungusian, an Altaic language. The accusative singular of animate nouns was the primary function of *\*m* in Indo-European, while Dravidian and the Altaic languages retained the original significance of the Nostratic formative, namely, that of the marked direct object.

8(E). According to Illič-Svityč (1971:214), the Nostratic particle *\*-da* was not a regular locative suffix, although Serebrennikov correctly presents Illič-Svityč's view on its original function. This functionally undifferentiated particle became a locative-ablative particle in Dravidian, a locative-ablative suffix in the Altaic languages, and an ablative suffix in locative adverbs in Uralic languages, which reveal its original significance of spatial deixis, cf. Erzya *aldo*, Hung. *aló-l* 'from below', South Hanty *to-l* 'from there', and an ablative marker in Indo-European (e.g., Skt. *ma-t* 'from me', and one can assume the presence of an originally undifferentiated locative-ablative particle in early Indo-European). This was the source of a spatial deictic particle in Hamito-Semitic: note directive *d* in Berber, e.g., *awi d* 'bring here', but cf. Tuareg *i-rvel d esen* 'he ran away from them'. Reflexes as a locative-lative particle are also found in Kartvelian. As we see it, the original absence of an exact semantic differentiation is found in all proto-languages except Indo-European. Serebrennikov's criticism is, therefore, unfounded; actually, he contradicts himself when he says that an ablative or lative meaning is not derivable from a locative, but then proceeds to give examples which, while they do not demonstrate such a course of semantic derivation, show the coexistence of these significances in one formative: namely, the ablative-locative in Turkic; the locative-lative in Mongolian--examples that reveal both the semantic and functional antiquity of the formative in question and that thus clearly corroborate Illič-Svityč's thesis.

9(E). However, Illič-Svityč never argued this.

10(E). Nevertheless, the preservation of several types

of collectivity might be one of the many archaisms of Uralic languages rooted in Nostratic.

11(E). This is not a point of contention, cf. fn. 10(E) above.

12(E). Serebrennikov's skepticism seems unfounded. Indo-European languages display archaic usages of this particle that could appear both before and after the verb, see Illič-Svityč (1971:250). This Nostratic particle is reflected in several Uralic languages, as well as in Dravidian, Indo-European, Kartvelian, and Hamito-Semitic: that is, in all Nostratic languages except Altaic, and this signals a high degree of preservation.

13(E). Serebrennikov does not appear to have been very attentive when reading this entry in Illič-Svityč's dictionary; nine lines in Illič-Svityč (1971:250-1) are devoted to Dravidian correspondences. As far as the absence of this particle in Altaic is concerned, this by no means speaks conclusively against its Nostratic origin, cf. fn. 12(E) above.

14(E). Here again Serebrennikov incorrectly estimates the age of the Nostratic proto-language (see fn. 2[E] above), and his contentions are thus fatuous.

15(E). This statement is totally unsubstantiated, see Ivanov's reviews of Illič-Svityč's work in this volume (pp. 1) in which he comments on the depth and persuasive nature of Illič-Svityč's comparisons. Serebrennikov's statement appears to be strictly *ad hominem* and perhaps derive from a moment of personal humiliation when, in 1964 at a conference of Soviet linguists, the then unknown Illič-Svityč spent some thirty minutes painstakingly listing Serebrennikov's errors in his studies of Uralic languages, his supposed specialty.

16(E). Note that, in what follows, Serebrennikov consistently contradicts this statement.

17(E). Unless Finnish is not a Uralic language, then Serebrennikov contradicts himself in this and the previous sentence, where, in fact, he actually acknowledges Illič-Svityč's studies for the inclusion of the Uralic languages.

18(E). For one thing, this is an example of the role Nostratic can play in understanding the high degree of antiquity of Finnish (that is, Uralic) vocalism. The next two sentences are the result of a misunderstanding, see fns. 2(E) and 14(E).

19(E). The reconstruction of a fully developed system of laryngeals, pharyngeals, and glottal stops for Nostratic is based on the West Nostratic languages, primarily Hamito-Semitic, see the tables in Illič-Svityč (1977:149) and the Foreword to this collection. Nevertheless, Uralic also has traces of Nostratic laryngeals, see Ivanov's article (pp. 14) and second review (pp. 57) in this collection.

20(E). See fns. 16(E), 17(E), and 18(E).

21(E). See fns. 2(E), 14(E), and 18(E).

22(E). This is an indication of the retention of an archaic stage of vocalism in both Turkic and Mongolian.

23(E). See fns. 2(E), 14(E), 18(E), and 22(E).

24(E). Data from Altaic and the reconstructions posited for Nostratic indicate that the correspondences cited by Serebrennikov are incorrect.

25(E). However, recent work by Voronin and others have shown that many onomatopoetic and sound symbolic correspondences are exceedingly stable. Independent work on Amerindian languages by the editors has led to the same conclusion.

26(E). As many of the above editorial notes show, this statement is without foundation.

27(E). The etymologies for these examples are implausible.

28(E). Virtually all of these correspondences are solidly grounded by Illič-Svityč (1971:Entries # 19, 131, 176). With respect to Finn. *käsi* 'hand' < Ural. *\*kāte-* < Nostr. *\*gāt'i* (Entry #80), Illič-Svityč cites the Uralic root with a question mark, but points out that Nostr. *\*t'* before *\*i* seemed to represent a strongly palatalized allophone of the type *T/S*, hence IE *\*s* (in IE *\*ġhes-* which supplied the inputs for the eastern dialects of Indo-European, e.g., Anatolian *\*ġhes-or-*, Indo-Iranian *\*ġhes-to-*, cf. Alb. *\*ġher-*, Gk. *\*ġhes-r-*, Arm. *\*ġhesr-m*, Toch. *\*ġher-s-*, whereas western dialects show a variety of secondary replacements and alterations, e.g., Italic *\*mān-*, Balto-Slavic *\*wronk-ā-*, Gmc. *\*kont-* > *\*hand-*, while Celtic employs, in part, the term for 'palm', *\*pāl-mā-*). Drav. *\*c* in *kac-id* derives from both Nostratic dental affricates and sibilants. Hamito-Semitic *\*c'* (in, for example, Cush. *\*kc'* 'hand') can be derived either from *\*kt'*, *\*kt'*, or *\*kš*. As for Buryat *gar-* 'go out', as well as 'protrude', Illič-Svityč explains this semantic development in Mongolian languages as deriving from an original meaning 'sharp branch, little

tree', cf. Tung. \**gara* in Manchu *gargan* 'branch', Naney *gara* id., Evenki *gara* 'twig, dry branch', Orok *garo* 'stickleback'. These correspondences have long been proposed and accepted by Altaists, and Illič-Svityč reconstructs the Altaic root \**gara* 'spike, branch, conifer' with which he compares Uralic \**kara* 'thorn, branch, conifer' Drav. *kar(a)* 'thorn, spike', IE \**ǵher(h)-* 'thorn, branch, spike' and reconstructs, altogether plausibly, a Nostratic protoform \**gara* 'thorny branch, thorn', see Illič-Svityč (1971:226, Entry #78).

29(E). The Nostratic character of Hamito-Semitic (also Semito-Hamitic) was substantiated by Illič-Svityč and, quite independently, by A. Dolgopolsky, in the early 1960s, see Ivanov's article and reviews in this collection. For an introduction to Hamito-Semitic/Semito-Hamitic, see the survey by Diakonoff (1965), now regarded as a classic.

30(E). It has always been clear that Semitic roots, as is the case for roots in any natural language, are based on real phonetic sequences that contain both consonants and vowels. Recent work by Dolgopolsky, Diakonoff, and others permit the reconstruction of concrete vowels, though without, of course, the details of surface phonetics, which must have been present in both Semitic and underlying Hamito-Semitic roots.

31(E). It is not altogether clear precisely why one is supposed to compare Finnish directly with Kartvelian, or, for that matter, Semito-Hamitic. It would be much more interesting and useful to compare reconstructed proto-languages: Uralic (from which Finnish ultimately derives) and Kartvelian, Indo-European, Hamito-Semitic, Dravidian, Mongolian, Turkic, Tungusian.

32(E). We doubt whether Illič-Svityč would have compared languages in this way, see the previous note. To see how Illič-Svityč interpreted possible correspondences between Uralic and Kartvelian, the object of Serebrennikov's rather biting criticism, see Illič-Svityč (1977: Entries #75, 131, 174, and 35, 166 respectively). In all of these cases, formal (= phonological/morphological) and semantic correspondences are correctly inferred, and Uralic-Kartvelian sets find cogent support in corresponding cognates in the other Nostratic languages.

33(E). Here, Serebrennikov correctly highlights some aspects of Illič-Svityč's methodology. We note, for example,

that, as different languages preserve different protolinguistic archaisms, the task for those reconstructing the proto-languages in question is to reveal these archaisms correctly and to reconstruct the phonological system of the proto-language on this basis. Since the East Nostratic languages, particularly Uralic, are seen as conserving the original Nostratic vocalism, it follows that the Proto-Nostratic vowel system is based in large measure on Proto-Uralic vocalism. On the other hand, two of the West Nostratic languages, Hamito-Semitic and Kartvelian, are deemed conservative as far as the obstruents are concerned, and the Proto-Nostratic obstruents are reconstructed mostly on the basis of Proto-Hamito-Semitic and Proto-Kartvelian. With respect to the Nostratic sonorants, they are reconstructed on the basis of the East Nostratic languages, as these languages are viewed as having preserved distinctions between different kinds of *n*-sonorants (*n*, *n̂*, *ñ*), *l*-sonorants (*l*, *l̂*, *l̃*), and *r*-sonorants (*r*, *r̂*). These distinctions were almost entirely lost in the Western Nostratic languages.

34(E). See the Foreword, pp. xxxvi.

35(E). The explanation for this is rather simple: Turkic *q* represents a guttural [k] before *a*, and it corresponds rather frequently to Semitic *q* which stands for [k'].

36(E). However, it is not the task of this kind of dictionary to answer glottogonic questions.

37(E). We hope that our comments--and they could have been expanded at length, to the present article, as well as the bulk of the present collection--make it abundantly clear that this conclusion by Serebrennikov and other like-minded linguists, who fail to recognize the possibility of discovering remote relationships between and among phyla, must now be dismissed as impossible or unsubstantiated, particularly so in light of recent advances in modern comparative theory. Here, too, one might point out that current work in universals by the UNITYP group in Cologne under the direction of Hansjakob Seiler is readily accessible for diachronic extensions that make the discovery of distant relationships, both typological and phylogenetic, a distinct possibility. Note, for example, the notions of implicational hierarchies, continua, inverse relationships between degrees of markedness and grammaticality, and other forms of scalar ranking employed

by UNITYP that permit comparison of seemingly diverse strategies and, thence, inferences about genetic affiliation. For example, case and classifier strategies, formerly analyzed independently of one another, may now be ranged on a continuum; see, for example, Markey (1985).

**On the Reconstruction of Proto-Indo-European Stops  
Glottalized Stops in Indo-European  
T. V. Gamkrelidze and V. V. Ivanov**

For late Proto-Indo-European the system of stops has been traditionally reconstructed as three sets with four points of articulation.<sup>1</sup> The series contain phonemes with identical phonetic features with respect to point of articulation, while sets are comprised of phonemes with the same mode of articulation. This means that sets unite homogeneous, but heterorganic phonemes.

The three sets of Indo-European occlusives are traditionally characterized as *voiced (mediae)*, *aspirated (mediae aspiratae)*, and *voiceless (tenuae)*; points of articulation series are characterized as *labial*, *dental*, *velar*, and *labiovelar*, see Table 1:

**THE TRADITIONAL (CLASSICAL) SYSTEM OF IE STOPS**

	<u>Sets</u>	<u>I voiced</u>	<u>II voiced aspirated</u>	<u>III voiceless</u>
Series	labial	(b)	bh	p
	dental	d	dh	t
	velar	g	gh	k
	labiovelar	g <sup>w</sup>	g <sup>w</sup> h	k <sup>w</sup>

In this traditionally reconstructed system, the disproportionate "filling" of the labial series of the voiced set attracts our immediate attention. Previously, Holger Pedersen noticed that, in such a system, a bilabial voiced phoneme, \*b, is absent in Indo-European; on the other hand, there are numerous root morphemes with g and d, cf. Pedersen (1951). All the forms employed by previous scholars to illustrate the presence of Indo-European \*b were dismissed by Pedersen as unconvincing. Specifically, Pedersen shows that, in the traditional set of stems, Skt. *balam* 'strength'; Gk. *βελίον* 'better'; OCS *bol'e* 'more'—only the Sanskrit and Slavic forms are cognates, and this does not permit the reconstruction of an IE \*b-.

Practically speaking, there are no forms that contain an incontestable instance of an IE *\*h*. Non-initial IE *\*-b-* has been recognized in but two forms, so Szemerényi (1970): Goth. *diups* : Lith. *dūgnas* : OCS *dŭno*, *dŭbri* : Gallic *Dubno-rīx* 'world king' : OIr. *domun* 'world' : Toch. A *tsopats*. But this set surely covers only one group of Indo-European dialects: namely, the Western group; and the primordial nature of the phoneme has been decided on the basis of Germanic alone. It is not impossible that this *-p-* in Gothic (and other Germanic languages) resulted from *b* < *\*bh* by "combinatorial" devoicing. Cf. Gk. βυδδῶς, (which, incidentally, is similar to the above set)—presupposes doublets *\*budh-/bhudh-* with possible metathesis of originally voiced phonemes in Gk. πυδμήν 'depth', see Chantraine (1968:201). As a result, neither of the above examples provides a secure basis for the reconstruction of a non-initial *\*-b-* in Indo-European. Another example, illustrating a possible IE *\*-b-* is as follows: ONw. *slapa* 'flabby' : OCS *slabu* 'weak' : Lith. *slobstū*, *slōb̃ti* 'become weak': usually compared with Lat. *labor*, *lapsus* 'to glide, slip; see appropriate entry in Pokorny (1959). Here, we are dealing with an areal delimitation of correspondences that presupposes later origin for these forms and, therefore, not a product of the Common Indo-European period. But even if we assume the existence of a PIE *\*b* in the above forms, then the numerical incongruity of forms with *\*b* on the one hand vs. those with *\*d* and *\*g* on the other hand is striking. According to our data, PIE forms containing *\*d* and *\*g* number over 250 each.<sup>2</sup>

Pedersen was the first to suggest a reinterpretation of this set of stops since *\*h* is lacking (or appears to be exceedingly rare) in PIE words; stops traditionally reconstructed as voiced should rather be considered voiceless, for, according to Pedersen, a language that lost *b*, but retained *d* and *g*, could hardly exist. On the other hand, there are numerous languages that have lost *p* but retained *t* and *k*. According to Pedersen, this original system had been transformed during the Common Indo-European period into the system that is traditionally reconstructed. As a typological parallel, Pedersen cites the development of voiceless and voiced stops in Eastern and Western Armenian languages, see Pedersen (1951).

It seems that such a process of transformation of an early Indo-European system as reconstructed by Pedersen into the system that is traditionally postulated was caused by difficulties encountered in the development of historically attested consonant systems derived from the Indo-European system proposed by Pedersen. The traditionally reconstructed system, which contradicts (as now becomes clear) the data of synchronic typology (and Pedersen was the first to notice this), has the quality of diachronic deductiveness; out of this system one can, easily and without discrepancies, deduce (by accepting typologically verifiable transformations—that is, transformations that are corroborated in the histories of many attested languages) all the systems of the historically attested Indo-European languages.<sup>3</sup>

The traditional picture of Indo-European consonantism was established at the outset of Indo-European comparative studies; and this picture coincided, by and large, with the consonant system of those Indo-European languages with the oldest literary records—languages such as Sanskrit and, in part, Greek and Latin. As these languages were often particularly prestigious for Indo-European grammar, they practically determined the configurations of the reconstructed system.

The systems of those languages that differed from that of the prestigious languages were considered the result of changes in the original system that was, in general, identically tailored to those systems in languages with the oldest textual traditions. Just so, Grimm, who, following Rask, was postulating the series of correspondences between Germanic and the Classical languages, considered the state of affairs in the Germanic languages as resulting from a transformation, a shift, of the original phonemes. This point of view—passed down from one generation of Indo-Europeanists to the next—is the prevalent view yet today in comparative Indo-European grammar.<sup>4</sup>

— — — — —

Now, as far as Indo-European vocalism is concerned, the justifiability of the traditional interpretations was evidenced fairly early on. The vocalism of Sanskrit turned

out to be secondary in relation to that found in other Indo-European languages. However, as far as the consonantism of Indo-European is concerned, few scholars doubted the Indo-European character of the Sanskrit system of stops, as well as (at least in part) the system of Greek and Latin.

The traditionally reconstructed system of Indo-European stops, though it displays the aura of diachronic deductionism, does not correspond to the data of synchronic typology. This incongruity is seen, first of all, in the lack (or, at the very least, great infrequency) of the voiced labial phoneme *\*b* of the first set.

According to the latest data on synchronic typologies, in consonant systems with the opposition voiced vs. voiceless, velar *g* is marked, while labial *b* is unmarked in the set of voiced stops, see Greenberg (1966, 1970), Hamp (1970), Melikišvili (1971, 1976), and Campbell (1973:44-6). As for the set of voiceless stops, here, in contrast, labial *p* is marked and velar *k* is unmarked. The frequencies of these corresponding phonemes (as well as empty cells, *cas vides*) are distributed in terms of these correlations. The marked member of the correlation is generally characterized by a lower frequency of occurrence than the unmarked member. In several systems, the low frequency of the marked member (*g*, *p*) can be identified with zero, that is, a *cas vide*, at the corresponding locus in the phonological system.<sup>5</sup>

Similar markedness correlations can be observed between different sets of stops in the subsystems of voiceless consonants. The most marked set in the subsystem of voiceless stops is that of the glottalized stops, which is more marked than the set of aspirated stops.<sup>6</sup> As for this latter set, it is more intensively marked than the set of "plain" voiceless stops. Thus, there is a hierarchy of markedness that can be represented as follows: plain--aspirated--glottalized voiceless stops, see Greenberg (1966). In line with this, the most intensively marked member in the sets discussed above is *p'* in the glottalized set, a state of affairs that results in its exceptional rarity, or, for that matter, its total absence in comparable systems in many languages that display glottalized consonants, such as, for example, some North Caucasian languages, numerous African languages, as well as Amerindian

languages in which  $p'$  is lacking and where, instead, we have to enter a *cas vide*, an empty slot, in the system, see Greenberg (1970), Hamp (1970).

The next incongruity between the system of Indo-European stops as traditionally reconstructed and the data marshalled from synchronic typology is the lack of a set of voiceless aspirated stops in a system in which there is simultaneously a set of voiced aspirated stops, an incongruity noted by Roman Jakobson, (1958:17-25). There are no attested languages in which voiceless aspirated stops are absent but in which voiced stops are present. In this respect, the system of Indo-European stops as traditionally reconstructed crassly contradicts the data of synchronic typology.

These incongruities necessitate a reconstruction of the traditional system in order to adjust it to facts obtained from synchronic typology. Such a reinterpretation of the Indo-European system of stops has to be performed in line with a process of diachronic deduction; and this presupposes a non-contradictory, typologically (that is, empirically) verifiable deduction from all historically attested structures extrapolated from the postulated system.

Of particular relevance for this reinterpretation is a redefinition of the distinctive features participating in the oppositions between and among the three sets of Indo-European stops, a redefinition keyed to both synchronic and diachronic typological data.

Let us now discuss each of the sets of the traditionally reconstructed system consecutively in terms of their mutual relationships.

The incongruity between the traditional system and the typological facts as mentioned by Jakobson (i.e., the simultaneous absence of voiceless and presence of voiced aspirated stops) can be eliminated if one reinterprets Set III as voiceless aspirated stops, a reinterpretation that accords fully with reflexes of this set in several of the historically attested Indo-European languages. In this manner, voiceless aspirated stops (Set III) now appear in the system where voiced aspirated stops (Set II) are simultaneously present, so that the traditional system now corresponds fully with the facts of synchronic typology, whereby deductions inherent in concrete, historically

attested systems extrapolated from it can be taken into consideration.

Of course, such a reinterpretation of Set III poses a question about its relationship with the other sets of the system, particularly Set I, which also has to be reinterpreted, as Set III receives a new interpretation. On the other hand, the reinterpretation of Set I is determined by its internal peculiarities, as well as the typological facts.

Set I, traditionally established as a series of voiced stops, must now be interpreted as a set of "non-voiced" stops because of the defectiveness of the labial component.

As pointed out above, the labial component is in fact defective (viz., either lacking or infrequently apparent systemically or textually) and therefore marked in the sets of non-voiced stops. As Joseph Greenberg (1970) has shown, the most highly marked among the sets is that of the glottalized stops, where *p'* is either entirely lacking or exceedingly rare. This universal characteristic of the glottalized set immediately poses a question about the possible phonological nature of Set I in Indo-European as a set of stops with a defective labial component. It is interesting to note that Set I is less frequent than Set II, to say nothing of Set III. According to Jucquois (1966), the general frequency of the members of Sets I-III in root morphemes is: Set I = 6.2%; Set II = 8.9%; Set III = 17.7%. This corresponds to absolute figures for the phonemic incidence of these sets in Indo-European stems as recorded in Pokorny's dictionary (our statistics).

These statistical correlations for the implementation of Indo-European "phonemes" alone cast doubt on their traditional interpretation as being voiced, voiced aspirated, and voiceless respectively. With respect to the universal rule of correlation between marked and unmarked phonemes, the set of voiced aspirated stops, which is marked in relation to the set of "plain" voiced stops, simply cannot be the more frequent set. If one interprets Set I as the set of glottalized stops, then one can see that the frequencies of these three sets are in full accordance with the typologically validated principle about the frequencies of glottalized stops in their relationships to other non-voiced sets in the phonological system, in particular to those of voiceless aspirated and plain voiceless stops.

The reinterpretation of Set III as that of voiceless aspirated stops is, first of all, based on typological considerations, as well as the composition of this set as reflected in historically attested languages, but is inconclusive for an interpretation of the phonemes of Set I (with a defective labial component): phonemes that are generally typologically specified as non-voiced, exactly as a set of glottalized stops, and not, say, as plain voiceless stops, to say nothing of voiceless aspirated stops. These two latter interpretations (plain voiceless, voiceless aspirated) are necessarily excluded for the simple reason that, in the subsystem of non-voiced stops, the set of glottalized stops is the most highly marked in comparison with the aspirated and plain voiceless sets.

Thus the defectiveness of the labial phoneme of Set I leads to the interpretation of this set as one of glottalized (ejective or "abruptive") phonemes, given the proviso that Set III is interpreted as a set of voiceless aspirated stops and that the above-mentioned data on statistical frequency are taken into consideration. Recast in such an interpretation, the Indo-European system of stops would now appear as follows:

Table 2

I	II	III
(p')	b <sup>h</sup>	p <sup>h</sup>
t'	d <sup>h</sup>	t <sup>h</sup>
K'	G <sup>h</sup>	K <sup>h</sup>

A system of Indo-European stops postulated in this manner corresponds fully to the facts of synchronic typology: defectiveness of the glottalized labial phoneme, functionally the weakest component of the group of non-voiced stops; simultaneous presence of voiced and voiceless aspirated stops; correlation with statistical frequencies that obtain between sets of the system that mirror the rising degrees of markedness in the sequence "voiced aspirated - voice aspirated - glottalized." We proposed such a system of Indo-European stops in a brief

expose as early as 1972, see Gamkrelidze and Ivanov (1972, 1973).

We would like to point out that similar notions about the presence of originally glottalized stops (instead of the traditional voiced stops) were proposed somewhat later by both Paul Hopper (1973:141-66), and André Haudricourt (1975). Apparently, Hopper arrived at similar conclusions independently of us; but he simultaneously took into consideration, on the basis of analogous typological reasoning, our earlier studies on the typologies of consonantism in reconstructed systems. Cf., too, André Martinet's (1953:70) early thoughts on the interpretation of defective series of stops as being primordially glottalized phonemes in his study of Semitic consonants.

In our proposed system of Indo-European stops, aspiration is actually a phonologically irrelevant feature, since Sets II and III are opposed, not on the basis of aspiration, but by virtue of voicedness vs. voicelessness. The feature of aspiration is to be viewed as a (non-distinctive) phonetic accompaniment. In strictly phonological terms, one could characterize each of the series as: I glottalized, II voiced, and III voiceless; but the phonetic quality of aspiration is a significant constituent of Sets II and III, which explains their vicissitudes and ultimate reflexes in the historically attested languages. Such phonetic constituents play a special role in diachronic phonetic changes, and one is compelled to consider them along with what are properly phonological features when reconstructing the inventory of stops.

Such aspiration in Sets II and III is phonologically irrelevant, for the corresponding phonemes of these sets could also be represented as allophonic variants without aspiration. This means that an aspirated stop, as well as its unaspirated distributional variant, could function as variants of a corresponding phoneme. Accordingly, each phoneme of Sets II and III could be represented by both of its allophones, aspirated and non-aspirated, depending on their privileges of occurrence in the word.

When allophonic variants of the reconstructed phonemes are taken into consideration, the inventory of Indo-European stops can be represented as follows:

Table 3

I	II	III
(p')	b <sup>h</sup> /b	p <sup>h</sup> /p
t'	d <sup>h</sup> /d	t <sup>h</sup> /t
K'	G <sup>h</sup> /G	K <sup>h</sup> /K

When laid out in this fashion, the system of Indo-European stops, now rendered typologically probable with respect to the phonemic correlations in the various sets, can also be considered typologically real in general. Systems of this structural type (with oppositions between glottalized and non-glottalized and voiced and voiceless) are found in many historically attested languages. For example, systems of voiced aspirated vs. plain voiced stops are recorded from several modern Armenian dialects in which voiced aspirated and plain voiced phonemes appear as positionally determined variants of one and the same phonemic unit, see Allen (1950), Džank'an (1967:78-81). The existence of such systems is, in effect, itself a sound empirical corroboration of the credibility of our reconstruction of Indo-European stops as a system with aspirated and non-aspirated variants of certain phonemes.

Allophones must have appeared depending on their position in the "word" when phonemes either came into contact (CC) or were in distant position (CVC; CRVC, CVRC), where C = consonant, V = vowel, and R = any syllabic resonant. Contact sequencing means uninterrupted constellations of two or more phonemes, whereas distant sequencing means a sequence of two consonant phonemes separated by one vowel or more phonemes in the frame of a reconstructed radical. If the vowel disappears, the distant sequence becomes a contact sequence, that is, a form with zero vowel.

The basic phonotactic rules governing Indo-European stops within the framework of a given radical can be formulated in terms of prohibitions and restrictions imposed on combinations of phonemes or their specific features. A major restriction on the entire subsystem of stops is the prohibition against the combination of identical

phonemes in the Indo-European radical, a restriction which can be formulated as Rule 1:

RULE 1: Two stops with identical matrices of distinctive features are incompatible (i.e., cannot occur simultaneously) in the frame of any given root of the type CVC; that is, there are no roots of the type  $T_1ET_2$ , where  $T_1 = T_2$ .<sup>7</sup>

The first restriction on the compatibility of phonemes in the frame of an Indo-European root is the prohibition against a distant sequence of two phonemes from Set I, the series of glottalized stops.

RULE 2: Two glottalized stops are incompatible within the frame of a single radical of the type CVC-; that is, there are no roots of the type  $t'ek'$ - =  $*deg$ - in the traditional notation.

This restriction on combinations of two glottalized stops in Indo-European has numerous typological parallels in languages with glottalized consonants. Thus, in native Kartvelian words, two non-identical glottalized consonants do not occur in the frame of a single root. The same type of restriction also holds in Shuswap, a Salishan language of British Columbia, where, in roots of the type  $C_1VC_2$  (as well as in roots with the format  $C_1RVC_2$ ,  $C_1VRC_2$ ), no glottalized  $C_1$  can appear if  $C_2$  is also a glottalized consonant, see Kuipers (1974:23). This is also true of several other Amerindian languages such as Maya (Yukatan dialect) and Quechua, as well as Hausa (Chadic group), see Carenko (1972), Rowe (1950), Orr and Longacre (1968). In Hausa, two heterorganic glottalized consonants never co-occur in one and the same word.<sup>8</sup>

Restrictions on the combination of glottalized consonants in the frame of a word evince the articulatory peculiarities associated with glottalized consonants that tend not to appear together in a distant sequence. In general, these restrictions are lifted when both glottalized consonants are identical and homorganic. But such combinations in the frame of an Indo-European root are unacceptable, see Rule 1.

Rule 2 is important for the typological justification of the interpretation of Set 1 stops as being specifically glottalized and not voiced, as was the case in the traditional system. An interpretation of Set I as voiced stops, the traditional view, lacks typological justification, as we are faced with the absence of Indo-European roots of

the type *\*ged-*, *\*deg-*, etc.: that is, roots in which both stops are voiced, as noted by Meillet—an absence that has remained unexplained until now, see Meillet (1924), Lehmann (1952), and Jucquois (1966).

Reinterpretation of the Indo-European "voiced" stops as glottalized removes this difficulty and also explains the lack of glottalized stops in terms of highly general and typologically verifiable phonotactic rules.

There are no restrictions on other combinations of glottalized stops with the stops of Set III in both initial and final positions, i.e., in all theoretically possible combinations that might occasionally occur.

RULE 3. Glottalized stops can appear in combination with all phonemes of Set III, both in initial and final position, so that the following types of constellations are possible: T<sup>[h]</sup>ET<sup>[h]</sup>, T<sup>[h]</sup>ET<sup>ʰ</sup>.

One readily notes the lack of combinations of phonemes of Set II with those of Set III and vice versa in Indo-European roots. This means the lack of roots of the type D<sup>[h]</sup>ET<sup>[h]</sup>, T<sup>[h]</sup>ED<sup>[h]</sup>, that is, in the traditional format, roots of the type *\*bhet-*, *\*tebh-*.

RULE 4. Both non-glottalized stops in the frame of one and same root have to be either voiced or non-voiced, thereby only permitting roots of the type D<sup>[h]</sup>ED<sup>[h]</sup> or T<sup>[h]</sup>ET<sup>[h]</sup>.<sup>9</sup>

A specific characteristic of the phonemes in Set II (voiced) and III (non-voiced) is the presence of aspiration as a phonetic feature. Each phoneme could be represented by either of its allophones (aspirated or non-aspirated), depending on its concrete phonetic environment. The environments in which these phonemes could be represented by both of their allophones determine the distribution of these phonemes.

There are certain linguistic data that prescribe, more or less precisely, the positions in which the phonemes of Sets II and III can appear as either aspirated or non-aspirated allophones. The dominant allophone of these sets is the aspirated one—for it appears in most phonetically independent positions, positions that can be established for Indo-European archetypes on the basis of comparison of historically attested, mutually correlated forms. A corresponding non-aspirated form of these phonemes is determined by its concrete environment.

This entails a distributional analysis of these phonemes in those positions (among all possible positions) in which they present themselves as non-aspirated allophones. Such an analysis yields more specific results when the phonemes of Set II are investigated. This set has left rather clear traces in the historically attested languages, a fact that permits one to reconstruct the distribution of the phonemes of this set rather exactly in the word forms of the Indo-European proto-language.

One of the main principles that determines the behavior of Set II phonemes in Indo-European protoforms is as follows: two such phonemes in the frame of one root are always represented by two different allophones, aspirated and non-aspirated. The distributional peculiarities of these allophones can be formulated as a rule:

In a root with two stops from Set II in a distant sequence, one of the phonemes is always represented by its aspirated allophone.<sup>10</sup> This means that a root can contain only one aspirated consonant. If the initial stop is represented by an aspirate and a non-aspirated allophone, then the next stop is represented by an aspirate, and vice versa; an aspirated initial allophone predicts the representation of the next phoneme by its non-aspirated allophone.

This distribution of Indo-European phonemes of Set II can be clearly observed and reconstructed on the basis of Indo-Iranian and Greek. These dialects reflect the distributional model of the allophones of Set II phonemes where non-aspirated ones are found in the succeeding position in the word—before a vowel or sonorant, e.g., Skt. *bahūs* 'arm' ~ Gk. *πῆχυς* 'elbow' presupposes an IE *\*bag<sup>h</sup>u-s* with non-aspirated initial and aspirated non-initial derived from IE */b<sup>h</sup>/* and */G<sup>h</sup>/* respectively with devoicing in Greek; Skt. *badhnāti*, later *bandhāti* 'he binds', Skt. *bandhus* 'relationship, bind, tie, affinity, relative', Gk. *πενθερός* 'father-in-law' (< 'bond by relationship'): IE *[\*bend<sup>h</sup>-]*; Skt. *bahū-* 'dense, multiple', Gk. *παχύς* 'thick, dense': IE *[\*benG<sup>h</sup>-]*; Skt. *bodhati*, *bodhate* 'he awakens, he wakes up', Gk. *πυνθόμηναι*, *πυνθάνομαι* 'I recognize, notice, am awake': IE *[\*beud<sup>h</sup>-]*, *[\*bund<sup>h</sup>-]*; Skt. *budhnā-* 'soil', Gk. *ζέφρο* id.: IE *[\*bud<sup>h</sup>-]*; Skt. *dahati* 'he burns' (tr.), *nidagha-* 'heat, summer', Gk. *ζεῖλον* 'ashes': IE *[\*deG<sup>h</sup>-]*; Skt. *dehmi* 'I anoint, smear', Skt. *dehi* 'wall', Gk. *τεῖχος* 'wall': IE *[deiG<sup>h</sup>-]*.

These examples from Sanskrit, together with their Greek analogues, clearly reflect the distribution of aspirated and non-aspirated allophones of corresponding Indo-European phonemes.

A similar distributional pattern for the allophones of Set II stops can be ascertained where the initial is reduplicated. Here, the reduplicated form of the roots \*dhē- 'put'; \*Ghē- 'deliver', \*bher- 'carry', etc., are particularly enlightening. Corresponding reduplicated forms are represented by IE \*da-dhā-mi and Gk. ζῆθῆ-μι: IE [\*di-d<sup>h</sup>e-mi] 'I give'; Skt. jahā-ti 'he delivers', Gk. \*κῑχημι 2nd sg. κῑχεῖς: IE [\*gi-G<sup>h</sup>-ē-ti], Skt. bi-bhar-ti < IE [\*bi-b<sup>h</sup>er-ti] 'he carries', cf. Gk εἰ-πιφθῶ, see Mayrhofer (1962).

The similarity (bordering on identity) between Sanskrit and Greek verbal forms with reduplication allows us, in the light of recent notions on the unity of the Greek-Aryan verbal system, to posit a common areal Indo-European source for these forms that corroborates the above thesis about the Indo-European source for the proposed phonemic deaspiration or the independent appearance of deaspirated phonemes in the histories of individual languages in particular, Sanskrit and Greek--such is the way in which Grassmann's Law is usually interpreted by traditional comparatists, (so Grassman 1862)--but they should rather be regarded as reflecting a common rule of distribution of these phonemes in an areal domain of the proto-language. The data from other Indo-European languages, such as Italic, Germanic, etc., are in full accordance with such an interpretation of these correlations. If so, then Grassmann's Law takes on a new guise; it is necessarily seen as an alternation of aspirated and non-aspirated sounds at the allophonic level in the system of the Indo-European proto-language, and not as the result of a process of deaspiration that took place independently in Sanskrit and Greek. A similarity approaching an identity of processes in Sanskrit and Greek can now be explained by the common origin of these processes as initiated in Common Indo-European.

For the Indo-European proto-language, this process is to be reconstructed, as clearly follows from the above, as a correlation of allophones that later became a phonemic alternation as a result of the phonologization of the reflexes

of voiced aspirated and non-aspirated allophones in the historically attested languages. After the phonologization of these reflexes of the stops of Set II in Greek, the well-established law of distribution of Indo-European aspirated and unaspirated allophones now becomes a productive rule of alternation of aspirated and unaspirated voiceless phonemes within the framework of a radical. In this manner, a productive morphological alternation of the following type can be explained: Gk. ζαχ<sup>ρ</sup>ος : θρ<sup>ι</sup>ξ (< IE \*drigh-), ζαχ<sup>ρ</sup>ος θ<sup>ε</sup>ος δ<sup>α</sup>ων (< IE \*d<sup>h</sup>negh-).<sup>11</sup>

Therefore, Indo-European protoforms for the above shapes in Greek must be reconstructed not as \*b<sup>h</sup>āgh-, \*b<sup>h</sup>engh-, \*b<sup>h</sup>eudh-, \*b<sup>h</sup>udh-, \*d<sup>h</sup>egwh-, \*d<sup>h</sup>eigh-, etc., with subsequent deaspiration of the reflexes of Indo-European voiced stops \*b<sup>h</sup> and \*d<sup>h</sup> independently in Sanskrit and Greek; but they have to be reconstructed as \*bagh-, \*bengh-, \*budh-, \*degwh-, \*deigh-, etc., with a positionally determined opposition between aspirated and non-aspirated allophones of voiced phonemes of Set II.

A system of Indo-European stops postulated in this manner is best preserved in Germanic, Armenian, and, probably, Anatolian (Hittite, Luwian). In the systems of these languages, insignificant phonetic variations can be assumed to permit the tripartite system of Indo-European stops as set forth above to be transformed into systems corresponding to those in these languages. Specifically, the first (glottalized) set of stops was reflected as unvoiced in Germanic; in Proto-Germanic, however, this set was probably characterized by glottal articulation.<sup>12</sup>

The second (voiced) set of Indo-European stops, with both aspirated and unaspirated allophones, underwent spirantization in Germanic and a subsequent evolution into corresponding voiced spirants, later becoming voiced stops initially and voiced fricatives non-initially. The aspects of spirantization, typical for these phonemes in Indo-European, contributed, it seems, to spirantization of Set II; and it is well known that aspirates tend to become spirants.

One could propose that voiced stops appeared initially, while voiced fricatives appeared non-initially, see Prokosch (1939). It is not excluded that this distribution of stops and fricatives in the reflexes of phonemes in Set II

corresponds to the archaic order of distribution of aspirated and non-aspirated allophones of Set II in Proto-Indo-European.

Reflexes of Set III in Germanic are parallel to those of the phonemes of Set II. The aspirated allophones of Indo-European voiceless stops are reflected in Germanic as voiceless fricatives, whereas the combinatorily conditioned non-aspirated allophones are, accordingly, represented in Germanic by non-aspirated reflexes (after *s-* and after a voiceless stop). In this respect, the transformation of phonemes of Set III in Germanic corresponds exactly to the transformation of phonemes of Set II; in this way, a parallelism in the development of phonetically similar sets of Indo-European stops is reflected.

In those forms that reflect a phoneme of Set III in a position before a stressed vowel (and after an unstressed one), the fricative was voiced according to Verner's Law, as in Goth. *fadar*, OIc. *faðir*, where *ð* is the result of voicing of a voiceless phoneme. Verner's Law stipulates a merger of originally (voiceless) fricatives *f*, *θ*, *χ*, *χ<sup>w</sup>* with their spirantized allophones, which, together with *b*, *d*, *g*, were united in the phonemic set *b/β*, *d/ð*, *g/ɣ*, and *g<sup>w</sup>/ɣ<sup>w</sup>*.<sup>14</sup> In this way, reflexes of the three original Indo-European sets of stops underwent a transformation in Germanic that resulted in the restructuring of the entire phonological system of Proto-Germanic.

Phonological transformations of the original Indo-European system, such as that outlined here for Germanic, that is, a transformation which eventuated the reshaping of the phonological systems of the historically attested Germanic dialects, are significantly different from those employed in traditional Indo-European studies to explain the genesis of Germanic consonantism, formalized as Grimm's Law, which covers, in particular, the first Common Germanic consonant shift.

The basic changes in the Germanic system of stops can actually be reduced to the phonetic process of spirantization of the allophones of the phonemes of Sets II and III. Though this phonetic process resulted in phonological restructuring of the Germanic system of stops (in comparison with that of Indo-European), one cannot characterize it as a wholesale "shift" of original

phonological sets in Germanic. The Germanic system conserves all the original features of the Indo-European system as far as [ $\pm$  voice and aspiration] are concerned. Thus, the Germanic phonological system is characterized by archaism and similarity to primordial phonological correlations as far as the features [ $\pm$  voice and aspiration] are concerned.

The Germanic system, together with that of Armenian and, probably, those of Anatolian and Common Tocharian (the proto-language of A and B), can be characterized as systems with a set of non-voiced stops (phonetically glottalized in some of these systems), originating from the Indo-European set of glottalized (non-voiced) stops. In this respect, these systems reflect some typical properties of Set I. As far as this feature is concerned, the above group of languages stands in opposition to another large group of Indo-European languages in which Indo-European glottalized stops (= Set I) have been transformed into voiced stops. To this group belong, among others, the Indo-Iranian languages, as well as Greek, Baltic, Slavic, Albanian, Celtic, and Italic. The glottalized Set I is here regularly reflected by plain voiced stops which presupposes voicing of originally glottalized phonemes.

This transformation of glottalized consonants into corresponding voiced stops can be observed in many languages and can be explained by the very nature of the articulation of glottalized consonants, i.e., with complete closure of the glottis. Glottal articulation is typical for voiced stops as well, an articulation that is characterized by constriction or closure of the vocal cords. The release of glottal closure in the course of articulating glottalized stops can be accompanied by a short vibration (release immediately following closure) of the vocal cords that is also typical for voiced stops according to the accompanying oral articulation. If the duration of the accompanying vibration is prolonged, then a corresponding voiced consonant (or voiced laryngealized consonant as in Hausa) can be produced—a consonant which, in general, has all the features of a glottalized stop. It is interesting that, when pronounced, glottalized sounds are closer to voiced than voiceless sounds on the continuum of throat operations.<sup>15</sup> Recent phonetic descriptions suggest that voiced sounds and those with glottal articulation

(laryngealized sounds) are closely related and belong to the same natural class. Phonetically, these sounds are much closer to each other than are sounds with glottal articulation to voiceless sounds, see Ladefoged (1971:19).

One can underscore the fact that, in many languages with quite diverse typologies, there are correspondences between glottalized stops and affricates vis-a-vis voiced stops and fricatives. Typical examples are Caucasian languages, as well as some Semitic languages. In the Veynach languages of the Northern Caucasus, Batsbian glottalized stops and affricates correspond, intervocalically and finally, to Chechen and Ingush voiced stops and fricatives, see Sommerfelt (1938).

In the Indo-European languages, that stop system which was most radically transformed in comparison with the original system is the Tocharian system where all three Indo-European sets have merged into one set, represented in writing by graphic signs for Indian voiceless and non-aspirated consonants. Next comes the system in the Baltic and Slavic languages, and then Celtic, where the phonetic character of the aspirated allophones of Set III is partially preserved. Next come the Italic and Greek systems in which there are many mutual innovations as far as the specific reflections of allophones of the phonemes of Set II are concerned. The Italic system preserves a voiced character of the unaspirated allophones of the phonemes of Set II, whereas in the Greek system both aspirated and non-aspirated allophones of the phonemes of this set have become devoiced.

Phonologically, even more significant transformations took place in the Indo-Iranian system with its symmetrical splitting of Sets II and III into four independent sets: voiced, voiced aspirated, voiceless, and voiceless aspirated. All of the above systems are united in a common group by virtue of how they reflect the Indo-European glottalized set; in all these systems, the glottalized set was devoiced and merged with the reflexes of Set II.

In this respect, the above group of languages is opposed to another group of Indo-European languages in which Set I stops preserved their originally voiceless character while Set III preserved its voiceless aspirated character. As members of this restricted group of archaic dialects, one can cite Anatolian, Germanic (with later spirantization

of Sets II and III), and Armenian—which preserved the original phonological correlations between the three sets of Indo-European stops to the largest degree—provided, that is, if Armenian possessed voiced aspirated and plain voiced stops as allophones of the common set of voiced phonemes. In this respect, Armenian can be considered the most archaic of the historically attested dialects.

The phonological transformations in most of the above Indo-European dialects can be described in terms of splits and mergers of the three original sets of Indo-European stops. Such processes were usual in those language systems (e.g., Indo-Iranian, Greek, and Italic) that have traditionally been considered the most archaic as far as the original consonantism is concerned. In contrast, those systems traditionally considered as the most dramatically restructured due to putative wholesale "shifts" now appear as the most archaic with respect to their consonantism; they are, in fact, closest to the original system of Indo-European stops. This implies that Grimm's Law as a description of the first Germanic consonant shift, the transformation of Indo-European consonantism in Germanic, is now inadequate with respect to the phonological processes postulated for Germanic itself. If one can speak of a "shift" in a specific, historically attested Indo-European system, then when evaluating the final results of this "shift" this assessment should apply, first and foremost, to such systems of Indo-European languages as those found in Indo-Iranian, Greek, Italic, and other languages that have traditionally been considered within Indo-European comparative studies as those systems that, generally speaking, reputedly reflect the original Indo-European system of consonantism.

## NOTES

**Ivanov/Gamkrelidze**

1) In a separate article (1980), we have discussed the differences in the local series of so-called "guttural" phonemes. In the present paper, in which we discuss the sets of Indo-European stops, we restrict ourselves to details

concerning the three local series: labial, dental, and velar, quite independently of how the velar series is to be defined further. This is why we employ unified symbols for both palatalized and velar gutturals in the present paper.

2) Statistical differences in frequencies between Jucquois' data (1966:61; 1971) and ours can, it appears, be explained by the fact that, in a way that differs from our approach, Jucquois counted consonant frequencies in root morphemes in Pokorny's dictionary, but considered as radicals only those shapes that fit Benveniste's definition of the root. Jucquois' data: *d* 83, *g* 70, *b* 31. It also seems that the frequential data obtained from Pokorny's dictionary do not reflect absolute correlations in Common Indo-European, for Pokorny also cites late lexemes that originated in the Indo-European dialects. This would explain why the frequency of *b* in Indo-European forms in Pokorny's dictionary is much higher than that for original Indo-European forms with *b*, some possible examples of which have been discussed here. Our statistics show *b* = 78 vs. Jucquois *b* = 31. Nevertheless, even these data, obtained as they are on the basis of all forms cited by Pokorny (both Common Indo-European and late, dialectal Indo-European) are enlightening as far as the correlation of frequencies is concerned: the frequencies of *d* and *g* are much higher than that for *b*.

3) This is probably one of the reasons why the system of Indo-European consonants as it is traditionally reconstructed was for so long viewed as the original one for those dialects that had been historically "fixed" with orthographic traditions, though when scrutinized more closely, this system too turned out to be contradictory, even with respect to synchronic typologies.

4(E). Interpretation of the Germanic consonants as the result of a shift (*\*th* < IE *\*t*, *t* < IE *\*d*, *\*θ* < *\*th* < IE *\*dh*) seems to be supported by several later shifts in the languages in this phylum, cf. also probable similar shifts in Hittite-Luwian and Armenian dialects. The Indo-European stops *\*T*, *\*D*, *\*Dh* are seen as deriving from Nostratic *\*T'*, *\*T*, *\*D* respectively, and this view excludes an hypothesis on the origin of Indo-European *\*b*, *\*d*, *\*g* from *\*p'*, *\*t'*, *\*k'*. Attempts to demonstrate a genetic relationship between IE *\*b*, *\*d*, *\*g* and Hamito-Semitic *\*p'*, *\*t'*, *\*k'* turned out to be disastrous,

see our remarks about Bomhard's work in this area in the Foreword.

It seems possible to take both the typological data marshalled by Jakobson and the putative shapes of Nostratic stops into account when reconstructing the Proto-Indo-European system. Accordingly, Germanic, Hittite-Luwian, and Armenian might indeed reflect an earlier stage of Indo-European, but not in the sense proposed by Gamkrelidze and Ivanov. One could, for example, assume that the most archaic structure of Proto-Indo-European stops is preserved by Armenian: Nostr. *\*T'*, *\*T*, *\*D* → IE: Arm. *\*Th*, *\*T*, *\*D*. This would be a development parallel to that in Altaic, cf. Turkic *\*Th*, *T*, *\*D* < Nostr. *\*T'*, *\*T*, *\*D*, and now see our remarks on this matter in the Foreword.

5) It is interesting to reflect on the contract between the improbability of lacking a *voiced* labial stop vs. the probability of lacking a *voiceless* labial stop. This contract corroborates Pedersen's contention that, historically, a phonological system could not expand after having lost the voiced labial stop (*b*), while languages that had lost the voiceless labial stop (*p*) could expand, both systemically and geographically. The significance of Pedersen's observation, particularly when contextualized by Greenberg's typological considerations, has been astutely recognized in Hamp's (1970) brief note. [(E). From the Nostratic viewpoint, IE *\*b* derives from Nostr. *\*p* and, albeit rarely, from Nostr. *\*p'*. Nostr. *\*p'* is usually reflected as *\*p* in Indo-European, whereas Nostr. *\*b* is always reflected as *\*bh* in Indo-European. From this, one might assume that, phonetically, this Nostr. (and early PIE) *\*p* already was [p:], cf. the state of affairs in the Northern Caucasian languages. Note that the remaining voiced stops (*\*d*: *\*g*, *\*ḡ*, *\*g<sup>w</sup>*) of Indo-European are derivable from Nostr. *\*t* and *\*k* respectively, but not from Nostr. *\*t'* and *\*k'/\*q'* (which are reflected as IE *\*t* and *\*k/\*ḱ/\*k<sup>w</sup>*). Cf. fn. 4 above.]

6) Here, the term "glottalized" is used in the narrow sense of "glottal occlusive" or "ejective," cf. Ladefoged (1971).

7) *T* is used to designate any stop; *T<sup>[h]</sup>* any voiceless (aspirated) stop; *T'* any glottalized stop; *D<sup>[h]</sup>* any voiced (aspirated) stop; so that each symbol represents the entire

set of corresponding stops. Therefore, a sequence of the type  $D^{[h^1]ED}$  or  $DED^{[h^1]}$  signals the fact that all units of a given set are possible in a certain radical, except for a combination of two identical phonemes, which is forbidden by application of Rule 1.

8) Cf. Parsson (1970:280). The origin of a specific tone in Lahu may be explained by a similar principle for the incompatibility of two glottal consonants, including glottal stops. (Lahu belongs to the Lolo subset of the Lolo-Birman sub-group of Birmanese languages.) See now Matisoff (1970).

9) Saussure was actually the first to discover this rule, cf. Meillet (1912:60), Benveniste (1935), Szemerényi (1972:143, fn. 56). On the basis of this rule, Kurytowicz presumed an opposition between the phonemes of Sets II and III, an opposition that distinguished not two distinctive features (aspiration and voice), but only one "phonemic feature"—all of which is rather close to our perception of the problem. However, Kurytowicz considered Set II neutral with respect to the opposition *voiced* : *voiceless*.

10) Here, one notes the following typological parallel: namely, the connection between the rule of incompatibility of aspirated stops in Quechua and the rule of incompatibility of glottalized stops, cf. Carenko (1972:102).

11) Sporadic forms in Greek with two aspirates, that is, the type  $\theta\upsilon\phi\lambda\acute{o}\varsigma$ , instead of regular  $\tau\upsilon\phi\lambda\acute{o}\varsigma$  'blind', can be regarded not as a reflection of the ancient distribution of two aspirates, but rather as a reflection of the rule for realization of aspiration in any stop in the sequence with a further generalization of aspiration for all segments of such a sequence. One should also recall that, subsequently, there was an assimilation of stops according to the feature of aspiration:  $(C...C^h) \rightarrow (C^h...C^h)$  or  $(C^h...C) \rightarrow (C^h...C^h)$ . This allows for forms with two aspirates as in Early Attic and Early Cretan  $\theta\acute{\epsilon}\theta\iota\varsigma$  instead of  $\theta\acute{\epsilon}\tau\iota\varsigma$ ;  $\theta\acute{\alpha}\nu\theta\alpha\iota\iota\varsigma$  instead of  $\Pi\alpha\mu\theta\alpha\iota\iota\varsigma$ , cf. also forms of the type  $\theta\upsilon\theta\acute{\epsilon}\nu$  'sacrificed' <  $\tau\upsilon\theta\acute{\epsilon}\nu$ , see Dressler (1975).

12) Traces of this glottal articulation in the first row in Germanic might be present in some prosodic features that are reflected in the later Germanic languages. One possible source for pharyngeal articulation in syllabic segments of the sorts reflected by Danish  $st\acute{o}d$ , the West Turkish glottal stop, and similar phenomena in Germanic

dialect words of the type *vā'tn* 'water' (cf. Katsnel'son 1966) is, it seems, the diffusion of glottalization over the entire syllable and concomitant transformation of this feature into a suprasegmental realization: IE *\*wot'-or/n* > *\*wa 't-n-*. The transformation of certain segmental features into suprasegmental features is both a well known and a widespread phenomenon in various languages.

13) It is interesting that Verner's Law applies only to voiceless fricatives and transforms them into voiced fricatives. That this law is not applied to Germanic voiceless stops is not readily comprehensible from a typological point of view. A probable explanation for this restriction may be provided by the character of the phonemic sets that have emerged in Germanic. For Germanic voiceless stops (which reflect Set I of the IE phonemes), the feature contrast *voiceless:voiced*, the basis for their opposition to the phonemes of Set II, was phonologically relevant; on the other hand, the fricatives of Set III were not in opposition to the independent row of fricatives with respect to the feature voicedness. The voicing of fricatives in this set in Germanic, due to the influence of the following stress accent, was, initially it seems, purely phonetic, and only later resulted in a specifically phonological redistribution of allophones. With this interpretation in mind, now cf. Normier (1977).

14) A phonological system emerges that is similar to that of Modern Spanish.

15) See Ladefoged (1971:16 ff.) and there note what is said about the correlation of glottalized, voiced, and laryngealized phonemes.

BIBLIOGRAPHY

- Adelung, Johann Christoph. 1781. "Über den Ursprung der Sprachen und den Bau der Wörter, besonders der deutschen. Ein Versuch. Leipzig: Breitkopf.
- Allen, William Sidney. 1950 (1951). Notes on the Phonetics of an Eastern Armenian Speaker. *Transactions of the Philological Society (Oxford)* 1950. 180-206.
- Andronov, M. S. 1982. *Iz istorii klassifikacii dravidijski jazykov. Teoretičeskie osnovy klassifikacii jazykov mira. Problemy rodstva.* Moscow: Nauka.
- Avram, A. 1978. L'opposition "implosif : explosif" et le problème des voyelles nasales du vieux slave. *Studia linguistica Alexandro Vasilii Filio Issatschenko a collegis amicisque oblata.* Ed. Henrik Birnbaum et al. Lisse: Peter de Ridder Press.
- Bacmeister, Hartwig Ludwig Christian. 1773. *Idea et desideria de colligendis linguarum specimenibus.* St. Petersburg (Petropoli): Russische kayserliche Akademie der Wissenschaften.
- Bader, Françoise. 1976. Le présent du verb 'être' en indo-européen. *Bulletin de la Société de Linguistique de Paris* 71.1.27-111.
- Bailey, H. W. 1957. A Problem of the Indo-Iranian Vocabulary. *Rocznik orientalistyczny* 21.59-69.
- Bańczerowski, Jerzy. 1972. Die Suche nach den uralischen Laryngalen. *Lingua Posnaniensis* 15.81-96.
- Benveniste, Emile. 1935. *Origines de la formation des noms en indo-européen.* Paris: Maisonneuve.
- Benzing, Johannes. 1956. *Die tungusischen Sprachen. Versuch einer vergleichenden Grammatik.* (Akademie der Wissenschaften und der Literatur. Abhandlungen der geistes- und sozialwissenschaftlichen Klasse, Jahrgang 1955, Nr. 11.). Wiesbaden: Steiner.
- Birnbaum, Henrik. 1977. Linguistic Reconstruction. Its Potentials and Limitations in New Perspective. *Journal of Indo-European Studies.* Monograph 2. Washington, D.C.: Institute for the Study of Man.
- Bomhard, Allan. 1984. *Towards Proto-Nostratic: A New Approach.* (Studies in the Theory and History of Linguistic Science. Series IV. Current Issues in Linguistic Theory. Vol. 27.). Amsterdam-Philadelphia: John Benjamins.

- Bulatova, R. 1977. Slav'anskij material v "Nostratičeskom slovarě" V. M. Illiča-Svityča. *Nostratičeskie jazyki* 74-9.
- Burrow, Thomas, and M. B. Emeneau. 1961. *A Dravidian Etymological Dictionary*. Oxford: Oxford University Press.
- Campbell, Lyle. 1973. On Glottalic Consonants. *International Journal of American Linguistics* 39.1.44-6.
- Carenko, E. I. 1972. O laringalizácii v jazyke kečua. *Voprosy jazyk o znaniya* 1.97-103.
- Čeyka, Mirek, and Lampret, Arnos; 1984. Nostratičnata ipoteza. Suvremenno sustojanie i perspektivi. *Contrastive Linguistics (Sofia)* 9.3.86-92.
- Chantraine, Pierre. 1968. *Dictionnaire étymologique de la langue grecque, histoire des mots*. I. Paris: Klincksieck.
- Cincius, Vera. 1984. Etimologii altajski leksem s anlautnymi gluximi pridyatelnymi smyčnymi gubnogubnym "p" i zadnejazyčnym "k". *Altajskie etimologii*. Leningrad: Nauka.
- Collinder, Bjorn. 1955. *Finno-Ugric Vocabulary. An Etymological Dictionary of the Uralic Languages*. Stockholm: Almqvist & Wiksell.
- . 1960. *Comparative Grammar of the Uralic Languages*. Stockholm: Almqvist & Wiksell.
- . 1966. Distant Linguistic Affinity. *Ancient Indo-European Dialects. Proceedings of the Conference on Indo-European Linguistics Held at the University of California, Los Angeles, April 25-27, 1963*. Ed. Henrik Birnbaum and Jaan Puhvel, 199-200. Berkeley & Los Angeles: University of California Press.
- Comrie, Bernard. 1981. *Language Universals and Linguistic Typology*. Oxford: Blackwell.
- Čop, Bojan. 1974a. *Indouralica*. 1. (Slovenska Akad. Znan. in Umetn. Classis II: Filologia et Litterae. Opera 30.1.). Ljubljana: Univerza v Ljubljani.
- . 1974b. Les changements paradigmatiques d'accentuation chez les thèmes nominaux en -e/o- indo-européens. *Linguistica (Ljubljana)* 14.39-53.
- Cowgill, Warren. 1975. More Evidence for Indo-Hittite: the Tense-Aspect System. *Proceedings of the Eleventh International Congress of Linguists*. Vol. 2. Ed. Luigi Heilmann, 557-70. Bologna: Società editrice il mulino.

- Cuny, Antoine. 1943. *Recherches sur le vocalisme, le consonantisme et la formation des racines en "nostratique". Ancêtre de l'indo-européen et du chamito-sémitique*. Lit-rarie d'amérique et d'orient). Paris: Maisonneuve.
- Diakonoff, Igor Miajlovič. 1963. *Urartskie pis'ma i dokumenty*. Moscow: Nauka.
- . 1965. *Semito-Hamitic Languages. An Essay in Classification*. Moscow: Nauka.
- . 1971. *Hurrisch und Urartäisch*. (Münchener Studien zur Sprachwissenschaft, Beiheft n.F.6.). Munich: R. Kitzinger.
- Dolgopolsky, A. B. 1964a. Metody rekonstrukcii obščeeindoevropejskogo jazyka i vne-indoevropejskie sopostavlenija. *Problemy sravnitel'noj grammatiki indoevropejski jazykov. Tezisy dokladov*. Moscow: Moscow State University.
- . 1964b. Gipoteza drevnejšego rodstva jazykov Severnoj Evrazii (problemy foneticeski sootvetstvija). VII *Meždunarodnyj kongress antropologičeski i etnografičeski nauk*. Moscow: Congress Publications.
- . 1969. Nostratičeskie osnovy s socetaniem šumnyx soglasny. *Ėtimologija* 1967. Moscow: Nauka. pp. 296-313.
- . 1973. *Sravnitel'no-istoričeskaja fonetika kušitski jazykov*. Moscow: Nauka.
- . 1974. O nostratičeskoj sisteme affrikat i sibil'antov: korni s fonemoj \*z. *Ėtimologija* 1972. Moscow: Nauka. Pp. 163-175.
- . forthcoming. On Personal Pronouns in the Nostratic Languages. *Gedenkschrift Björn Collinder*.
- Dondua, Karpez Darispanovič. 1975. Franz Bopp i kavkazskie jazyki. *Stat'ji po obščemu i kavkazskomu jazykoznaniju*. Moscow: Nauka.
- Dressler, Wolfgang Ulrich. 1975. Zur Rekonstruktion phonologischer Prozesse im Altgriechischen: Grassmanns Gesetz. *Bereiche der Slavistik. Festschrift zur Ehren von Josip Hamm*. Pp. 53-67. Vienna: Österreichische Akademie der Wissenschaften.
- Dybo, Vladimir Antonovič. 1978. Nostratičeskaja gipoteza (itogi i problemy). *Izvestija AN SSSR, ser. lit-ry i jazyka*, 1978, t. 37, No. 5.
- Džauk'an, Gevork Beglarovič. 1967. *Očerki po istorii dopis'mennogo perioda arm'anskogo jazyka*. Erevan. Nauka.

- Foucault, Michel. 1971. *The Order of Things. An Archaeology of the Human Sciences*. New York: Random House. Transl. of: *Les mots et les choses*. Paris: Gallimard, 1966.
- Fussman, Gérard. 1972. *Atlas linguistique des parlers dardes et kafirs. II. Commentaire*. (Publications de l'Ecole française d'Extrême-Orient, 86.). Paris: Maisonneuve.
- Gamkrelidze, Tamaz Valerianovič. 1959. (*Sibilant Correspondences and Several Problems of the Ancient Structure of the Kartvelian Languages*). Tbilisi: Nauka. (Original in Georgian).
- Gamkrelidze, Tamaz Valerianovič, Ivanov, V'ačeslav, Vsevolodovič,  
 — 1972. Lingvističeskaja tipologija i rekonstrukcija sistemy indoevropejskix smecnyx. *Working Papers of the Conference on the Comparative Grammar of the Indo-European Languages* (December 12-14, 1972). Moscow: Nauka.  
 — 1973. Sprachtypologie und die Rekonstruktion der gemeinindogermanischen Verschlüsse. Vorläufiger Bericht. *Phonetica* 27.150-6.
- Georgiev, Vladimir I. 1981. *Introduction to the History of the Indo-European Languages*. Sofia: Bulgarian Academy of Sciences. 3rd ed.
- Giginejšvili, B. K. 1977. *Sravnitel'naja fonetika dagestanskix jazykov*. Tbilisi: Tbilisskij universitet.
- Glossarium comparativum totius orbis*. 1787. St. Petersburg: Russische kayserliche Akademie der Wissenschaften. 2nd ed. St. Petersburg: Russische kayserliche Akademie der Wissenschaften, 1790-1791. Vols. 1-4.
- Grassmann, Hermann. 1862. Über die Aspiraten und ihr gleichzeitiges Vorhandensein im An- und Auslaute der Wurzeln. (*Kuhn's*) *Zeitschrift für vergleichende Sprachforschung auf dem Gebiete der indogermanischen Sprachen*. 12.81-138.
- Greenberg, Joseph H. 1952. Historical Linguistics and Unwritten Languages. *Anthropology Today. An Encyclopedic Inventory*. Ed. A. L. Kroeber, 265-87. Chicago: University of Chicago Press.  
 — 1957. *Essays in Linguistics*. Chicago: University of Chicago Press.  
 — 1966. *Language Universals with Special Reference to Feature Hierarchies*. The Hague-Paris: Mouton.

- . 1970. Some Generalizations Concerning Glottalic Consonants, Especially Implosives. *International Journal of American Linguistics* 36.2.123-45.
- Grillot, Françoise. 1978. Les affixes nominaux et les pronoms indépendants de la langue élamite. *Journal asiatique* 266.1-2.1-35.
- Güldenstädt, Johann Anton. 1787-1791. *Reisen durch Russland und im kaukasischen Gebirge. Auf Befehl der russischen kayserlichen Akademie der Wissenschaften*. Ed. P. S. Pallas. Vols. 1-2. St. Petersburg: Russische kayserliche Akademie der Wissenschaften.
- Haas, Volkert, and Gernot Wilhelm. 1974. Zum hurritischen Lexicon. II. *Orientalia (Pontificum Institutum Biblicum)* 43 (N. S.). 1.87-93.
- Hamp, Eric P. 1970. Maya-Chipaya and Typology of Labials. *Papers from the Sixth Regional Meeting (Chicago Linguistic Society)*. Pp. 20-2. Chicago: University of Chicago Department of Linguistics.
- Haudricourt, André. 1975. Les mutations consonantiques (occlusives) en indo-européen. *Mélanges linguistiques offerts à Emile Benveniste*. Louvain: Editions Peeters. Pp. 267-272.
- Haudry, Jean. 1981. *Les indo-européennes*. (Que sais-je?). Paris: Presses Universitaires de France.
- Hill, Jane H. 1969. Volitional and Non-Volitional Verbs in Cupeño. *Papers from Fifth Regional Meeting (Chicago Linguistic Society)*. Pp. 348-56. Chicago: University of Chicago Department of Linguistics.
- Hodge, Carlton T. 1974. Review of W. B. A. Lockwood, *A Panorama of Indo-European Languages*. *Language Sciences* 30.32-5.
- Hopper, Paul J. 1973. Glottalized and Murmured Occlusives in Indo-European. *Glossa* 7.2.141-66.
- . 1982. Areal Typology and the Early Indo-European Consonant System. *The Indo-Europeans in the Fourth and Third Millennia*. Ed. Edgar C. Polomé, 121-39. Ann Arbor: Karoma Publishers, Inc.
- Illic-Svityč, Vladislav Markovič. 1963. Drevneysiě indoevropejsko-semitskie jazykovye kontakty. *Problemy indoevropejskogo jazykoznanija*. Moscow: Nauka.
- . 1967. Materialy k sravnitel'nomu slovar'u nostratičeskix jazykov. *Étimologija* 1965. Pp. 307-373.
- . 1971. *Opyt sravnenija nostratičeskix jazykov*.

- Sravnitel'-nyj slovar' (b-k)*. Moscow: Nauka.
- 1976. *Opyt sravnenija nostraticheskix jazykov. Sravnitel'-nyj slovar' (l-z)*. Ukazateli. Moscow: Nauka.
- 1984. *Opyt sravnenija nostraticheskix jazykov. Sravnitel'-nyj slovar' (p-q)*. Ed. A. Dybo. (Fascicle 1 of Vol. 3). Moscow: Nauka.
- Ivanov, V'acheslav V. 1959. Toxarskie jazyki i ix znachenie dl'a sravnitel'no-istoricheskogo issledovanija indoevropejskix jazykov. *Tokharskie jazyki*. Moscow: Nauka.
- 1974. Iz etimologicheskix nabl'udenij nad baltijskoj leksikoy. 1. Latysh. *māsa* 'sestra' i i.-e. nazvanija 'sestry'. *Zeitschrift für Slawistik* 19 (Hft. 2). *Beiträge zur Baltistik*.
- 1977a. Review of W. P. Lehmann, *Proto-Indo-European Syntax*. *Izvestija AN SSSR, Ser. lit. i jaz.*, t. 36, No. 4.
- 1977b. The Relation between Different Grammatical Levels in the Linguistic Evolution. *Die Sprache* 23.1.20-4.
- 1978. O sledakh indo-evropejskix kompleksov enklitik v slav'anskom. *Studia linguistica Alexandro Vasilii Filio Issatschenko a collegis amicisque oblata*. Ed. Henrik Birnbaum et al. Lisse: Peter de Ridder Press.
- 1983. *Istorija slav'anskix i balkanskix nazvanij metallov*. Moscow: Nauka.
- Ivanov, V'acheslav V., and Gamkrelidze, Tamaz V. 1980. R'ad "gutturalnj" v indoeuropejskom. Problema jazykov centum i sat m i otaženie "guttural'nyx" v istoričeskix indoeuropejskix dialektax. *Voprosy jazykoznanija* 5.10-20, 6.13-22.
- Jakobson, Roman. 1958. Typological Studies and their Contribution to Historical Comparative Linguistics. *Proceedings of the Eighth International Congress of Linguists 1957*. Pp. 17-25. Oslo: Oslo University Press.
- Jasanoff, Jay H. The r-Endings of the Indo-European Middle. *Die Sprache* 23.2.159-70.
- Jucquois, Guy. 1966. La structure des racines en indo-européen envisagée d'un point de vue statistique. *Linguistic Research in Belgium*. Ed. Y. Lebrun, 57-68. Wetteren: Universa.
- 1970. Les postpositions due hittite et l'accentuation des preverbes en indo-européen. *Le Museon* 83.3/4.533-40.

- . 1971. La théorie de la racine en indo-européen. *La linguistique* 7.1.73-91.
- Kačnel'son, Solomon Davidovič. 1966. *Sravnitel'naja grammatika germanskix jazykov*. Moscow: Nauka.
- Klimov, Georgij Andreevič. 1969. Abxazskoadygskokartvel'skie leksičeskie paralleli. *Ėtimologija* 1967. Pp. 286-295.
- . 1977. *Tipologija jazykov aktivnogo stroja*. Moscow: Nauka.
- Krejnič, Eruxim Abramovič. 1958. *Jukagirskij jazyk*. Leningrad: Nauka.
- Krejnič, E. A. 1968. *Glagol ketskogo jazyka*. Leningrad: Nauka.
- Kuipers, Aert H. 1974. *The Shuswap Language. Grammar. Text. Dictionary*. The Hague-Paris: Mouton.
- Kurylowicz, Jerzy. 1973. Internal Reconstruction. *Current Trends in Linguistics. Vol. 11. Diachronic, Areal, and Typological Linguistics*. Pp. 63-92. The Hague-Paris: Mouton.
- Ladefoged, Peter. 1971. *Preliminaries to Linguistic Phonetics*. Chicago: University of Chicago Press.
- Lehmann, Winfred P. 1952. *Proto-Indo-European Phonology*. Austin: University of Texas Press.
- . 1974. *Proto-Indo-European Syntax*. Austin-London: University of Texas Press.
- Lévy, Paul. 1972. Grec *thálassa* 'mer' < Sanskrit *tātaka* 'étang' > Siamois *thal'é* 'mer'. *Langues et techniques. Nature et société. Tome I. Approche linguistique*. Pp. 295-300. Paris: Klincksieck.
- Lewy, E. L. 1961. *Kleine Schriften*. (Deutsche Akademie der Wissenschaften zu Berlin, Veröffentlichungen der sprachwissenschaftlichen Kommission, Bd. 1.). Berlin: Akademie: Verlag.
- Lyons, John. 1968. *Introduction to Theoretical Linguistics*. Cambridge: Cambridge University Press.
- Markey, T. L. 1983. Gmc. *\*baina-* 'bone' and Other Monstrosities. *NOWELE* 2.93-107.
- . 1985. The Case vs. Classifier Continuum. *Zeitschrift für Dialektologie und Linguistik* 52.
- McLendon, Sally A. 1978. Ergativity, Case and Transitivity in Eastern Pomo. *International Journal of American Linguistics* 44.1.1-9.
- Martinet, André. 1953. Remarques sur le consonantisme sémitique. *Bulletin de la Société de Linguistique de*

- Paris 41.1.67-78.
- 1978. Des jers slavs aux voyelles caduques du japonais. — *Studia linguistica Alexandro Vasilii Filio Issatschenko a collegis amicisque oblata*. Ed. Henrik Birnbaum et al., Lisse: Peter de Ridder Press.
- Matisoff, J. A. 1970. Glottal Dissimilation and the Lahu High-Rising Tone: a Tono-genetic Case-study. *International Journal of American Linguistics* 90.1.
- Matouš, Lubor. 1978. Zum System der Datierung in der Handelskolonie Kaněš. *Archivorientální* 46.3.217-31.
- Mayrhofer, Manfred. 1963. *Kurzgefasstes etymologisches Wörterbuch des Altindischen/A Concise Etymological Sanskrit Dictionary*. Heidelberg: Winter.
- Meillet, Antoine. 1913. À propos de avestique zrazā- ~~Mémoires de la Société de Linguistique de Paris~~ 18. 60-4.
- 1924. *Introduction a l'étude des langues indo-européennes*. Paris: Hachette. 6th ed.
- Melikišvili, Georgij Aleksandrovič. 1964. *Urartskij jazyk*. Moscow: Nauka. (German transl. by Karl Schrembek. With an Appendix by Mirjo Salvini. Ed. by A. Kammenhuber and M. Salvini. (Studia Pohl, 7.). Rome: Biblical Institute, 1971).
- 1971. Urartskie klinoobraznye nadpisi. *Vestnik drevnej istorii* 1971. Moscow: Nauka.
- Melikišvili, I. G. 1971. K izučeniju ierarxičeskix otnošenij edinic fonologičeskogo urovn'a. *Voprosy jazykoznanija* 3.
- 1976. (*Relationship of Markedness in Phonology*). — Tbilisi: Mecenereba. (Original in Georgian).
- Normier, R. 1977. Indogermanischer Konsonantismus, germanische "Lautverschiebung" und Vernersches Gesetz. (*Kuhns*) *Zeitschrift für vergleichende Sprachforschung auf dem Gebiete der indogermanischen Sprachen* 91.171-218.
- Nostratičeskie jazyki. 1977. *Nostratičeskie jazyki i nostratičeskoe jazykoznanie. Tezisy dokladov. Konferencija*. Moscow: Nauka.
- O'Hear, Anthony. 1980. *Karl Popper*. London: Routledge & Kegan Paul.
- Onians, R. B. 1954. *The Origins of European Thought about the Body, the Mind, the Soul, the World, Time and Fate. New Interpretations of Greek, Roman and*

- Kindred Evidence and of Some Basic Jewish and Christian Beliefs.* Cambridge. Cambridge University Press. 2nd ed.
- Orr, Carolyn, and Longacre, Robert E. 1968. Proto-Quechumaran. *Language* 44.3.528-55.
- Otten, Heinrich. 1973. *Eine althethitische Erzählung um die Stadt Zalpa.* (Studien zu den Boğazköy-Texten, Hft. 17.). Wiesbaden: Harrassowitz.
- Parsons, F. W. 1970. Is Hausa Really a Chadic Language? Some Problems of Comparative Phonology. *African Language Studies* 11.272-97.
- Pedersen, Holger. 1899 (1900). Albanisch und Armenisch. (Kuhns) *Zeitschrift für vergleichende Sprachforschung auf dem Gebiete der indogermanischen Sprachen* 36. 340-1.
- . 1903. Türkische Lautgesetze. *Zeitschrift der deutschen morgenlandischen Gesellschaft* 57.535-61.
- . 1935. Il problema delle parentele tra i grandi gruppi linguistici. *Atti del' III Congresso internazionale dei linguisti (Roma, 19-26 sett., 1933).* Pp. 328-33. Florence: Felice de Monnier.
- . 1938. *Hittitisch und die anderen indoeuropäischen Sprachen.* (Det Kgl. Danske Videnskabernes Selskab, Historisk-filologiske Meddelelser, 25.2.). Copenhagen: Munksgaard.
- . 1950. Eine tocharische Frage. *Jahrbuch für kleinasiatische Forschung* 1.1.103-4.
- . 1951. *Die gemeindoeuropäischen und die vorindoeuropäischen Verschlusslaute.* Det Kgl. danske Videnskabernes Selskab, Historisk-filologiske Meddelelser, 32.5.). Copenhagen: Munksgaard.
- Pil'sčikova, N. N. 1959. Mestoimenija jazykov nigeročadskoj gruppy. *Afrikanskij ètnografičeskij sbornik. III.* (Trudy Instituta ètnografii im. Mikluxe-Maklaja. Novaja serija, t. II.). Moscow-Leningrad: Nauka.
- Pisani, Vittore. 1959. *Saggi di linguistica storica; scritti scelti.* Torino: Rosenberg & Sellier.
- Pokorny, Julius. 1959. *Indogermanisches etymologisches Wörterhurch.* I. Bern: Francke Verlag.
- Polivanov, Evgenij Dmitreivič. 1928. *Vvedenie v jazykoznanie dl'a jazykovednyx vuzov.* 1. Moscow: Nauka. 1960.
- . 1960. *Predvaritel'noe soobščenie ob ètimologičeskom slo-vare japonskogo jazyka.* *Problemy vostokovedenija*

- Polomé, Edgar C., Ed. 1982. *The Indo-Europeans in the Fourth and Third Millennia*. Ann Arbor: Karoma Publishers, Inc.
- Prokosch, Eduard. 1939. *A Comparative Germanic Grammar*. (William Dwight Whitney Linguistic Series). Philadelphia: Linguistic Society of America.
- Ramstedt, Gustaf John. 1924. *Die Verneinung in den altaischen Sprachen*. (Mémoires de la Société finno-ougrienne, 52.). Helsinki: Druckerei der finnischen Literaturgesellschaft.
- . 1957. *Einführung in die altaische Sprachwissenschaft*. I. Lautlehre. Ed. and rev. Pentti Aalto. Helsinki: Suomalais-ugrilainen seura.
- Riedl, Rupert. 1979. *Biologie der Erkenntnis. Die stammesgeschichtlichen Grundlagen der Vernunft*. Berlin-Hamburg: Verlag Paul Parey.
- Rosenkranz, Bernhard. 1966. Zur indo-uralischen Frage. *Istituto orientale di Napoli. Annali, Sezione linguistica*, VII. Rome.
- Rowe, John Howland. 1950. Sound Patterns in Three Inca Dialects. *International Journal of American Linguistics* 16.3.137-48.
- Schmidt, Karl Horst. 1957. Die Komposition in gallischen Personennamen. *Zeitschrift für celtische Philologie* 26.3/4.33-301.
- Schuchardt, Hugo. 1925. *Der Individualismus in der Sprachforschung*. *Sitzungsberichte der Wiener Akademie* 202.4.1-21.
- Serres, Michel. 1982. *Hermes, Literature, Science, Philosophy*. Ed. Josué V. Harari and David F. Bell. Postface by Ilya Prigogine and Isabelle Stengers. Baltimore-London: The Johns Hopkins University Press.
- Setälä, Emil Nestor. 1915. *Zur Frage nach der Verwandtschaft der finnischugrischen und samojedischen Sprachen*. Helsinki: Finnisch-ugrische Gesellschaft.
- Shanmugam, S. V. 1971. *Dravidian Nouns. A Comparative Study*. (Department of Linguistics, Publication No. 25.). Annamalainagar: Annamalai University.
- Shevoroshkin, Vitalij. 1981. Dene-Caucasian Macrofamily, *Athapascan Quarterly* 2.1.
- Sommerfelt, Alf. 1938. Etudes comparatives sur le caucasi-

que du Nord.-Est. *Norsk tidskrift for sprogvidenskap* 9.115-43.

Speiser, Ephraim Avigdor. 1941. *Introduction to Hurrian*. (Annual of the American Schools of Oriental Research.

Starostin, Sergej A. 1982. *Praenisejskaja rekonstrukcija i vnešnie sv'azi enisejskix jazykov. Ketskij sbornik*. Leningrad: Nauka.

Sturtevant, Edgar H., and Hahn, E. Adelaide. 1951. *A Comparative Grammar of the Hittite Language*.

(William Dwight Whitney Language Series). New Haven: Yale University Press. 2nd ed. rev.

Swadesh, Morris. 1952. Lexico-statistic Dating of Pre-historic Ethnic Contacts with Special Reference to North American Indians and Eskimos. *Proceedings of the American Philosophical Society* 96.452-63.

—. 1955. Towards Greater Accuracy in Lexicostatistic Dating. *International Journal of American Linguistics* 21.2.121-37.

—. 1960. *American-Eurasian Linguistic Relations*. Unpublished MS.

—. 1962a. Archaic Doublets in Altaic. *American Studies in Altaic Linguistics* Ed. Nicholas Poppe, 293-330. (Indiana University Publications, Uralic and Altaic Studies, Vol. 13.). Bloomington: Indiana University Press and The Hague: Mouton & Co.

—. 1962a. Linguistic Relations Across the Bering Strait. *American Anthropologist* 64.6.1262-91.

Szemerényi, Oswald John Louis. 1970. *Einführung in die vergleichende Sprachwissenschaft*. Darmstadt: Wissenschaftliche Buchgesellschaft.

—. 1972. Comparative Linguistics. *Current Trends in Linguistics*. Vol. 9.1. *Linguistics in Western Europe*.

Ed. T. A. Sebeok, 119-95. The Hague-Paris: Mouton.

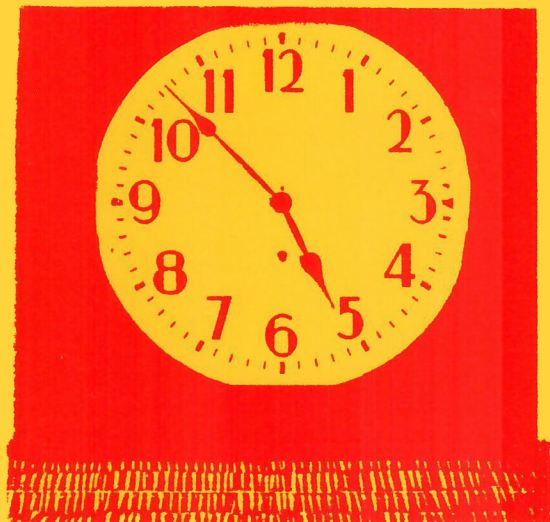
Tischler, Johann. 1983. *Hethitisches etymologisches Glossar. Teil 1. a-k*. (Innsbrucker Beiträge zur Sprachwissenschaft, Bd. 20.). Innsbruck: Institut für Sprachwissenschaft.

Trubetzkoy, N.S. 1929. *Notes sur les desinences du verbe des langues tchétochéno-lesghiennes (caucasiques orientales)*. Bulletin de la Société de Linguistique de Paris 29.153-71.

—. 1939. *Gedanken über das Indogermanenproblem*. *Acta Linguistica* 1.81-9.

- Tungusian*. 1975. *Sravnitel'nyj slovar' tunguso-mančžurskix jazykov. Materialy k etimologičeskomu slovar'u*. Vol. 1. A- . Leningrad: Nauka.
- Tungusian*. 1977. *Sravnitel'nyj slovar' tunguso-manc zurski jazykov. Materialy k etimologičeskomu slovar' u*. Vol. 2. O-#. Leningrad: Nauka.
- Wang, William S. -Y. 1976. Language Change. *Annals of the New York Academy of Sciences (Origins and Evolution of Language and Speech)*. 280.61-72.
- Watkins, Calvert. 1969. *Indogermanische Grammatik. Bd. III. Formenlehre. Erster Teil. Geschichte der indogermanischen Verbal flexion*. Heidelberg. Winter.
- Xelimskij, Evgenij Arnol'dovič. 1982. *Drevnejšie vengersko-samodijskie jazykovye paralleli*. Moscow: Nauka.
- Zadok, Ran. 1975. Iranian Names in Late Babylonian Documents. *Indo-Iranian Journal* 17.3/4.245-7.
- Zelenin, Dmitrij Konstantinovič. 1927. *Russische (ostslavische) Volkskunde*. Berlin-Leipzig: Walter de Gruyter.
- Zvelebil, Kamil. 1977. *A Sketch of Comparative Dravidian Morphology*. (Janua Linguarum, Series Practica, 180.). The Hague: Mouton.





---

This volume presents a collection of recent seminal papers on language change, relationship, and typology by prominent Soviet linguists. The central theme here is a discussion of the pros and cons of the Nostratic hypothesis, the radical view which has figured as a focal point for Soviet comparativists that a wide variety of seemingly unrelated and widely geographically separated languages --- ranging from Indo-European to Semito-Hamitic, from Finno-Ugric to Dravidian and the various languages of the Caucasus --- all stem from a single distant parent.

These papers are selected from works by such outstanding scholars as A. B. Dolgopolsky, B. A. Serebrennikov, V. V. Ivanov, and T. V. Gamkrelidze, all of whom were deeply indebted to the pioneering efforts of V. M. Illič-Svityč. Here, too, one will find the revolutionary paper on the development of the Indo-European consonant system by Gamkrelidze and Ivanov. The papers are prefaced by a critical intro by the editors and translators in an attempt to familiarize English-speaking readers with this important new sector of linguistic inquiry.



---

KAROMA PUBLISHERS, INC.  
ISBN: 0-89720-072-1 (Paperback Only)

---